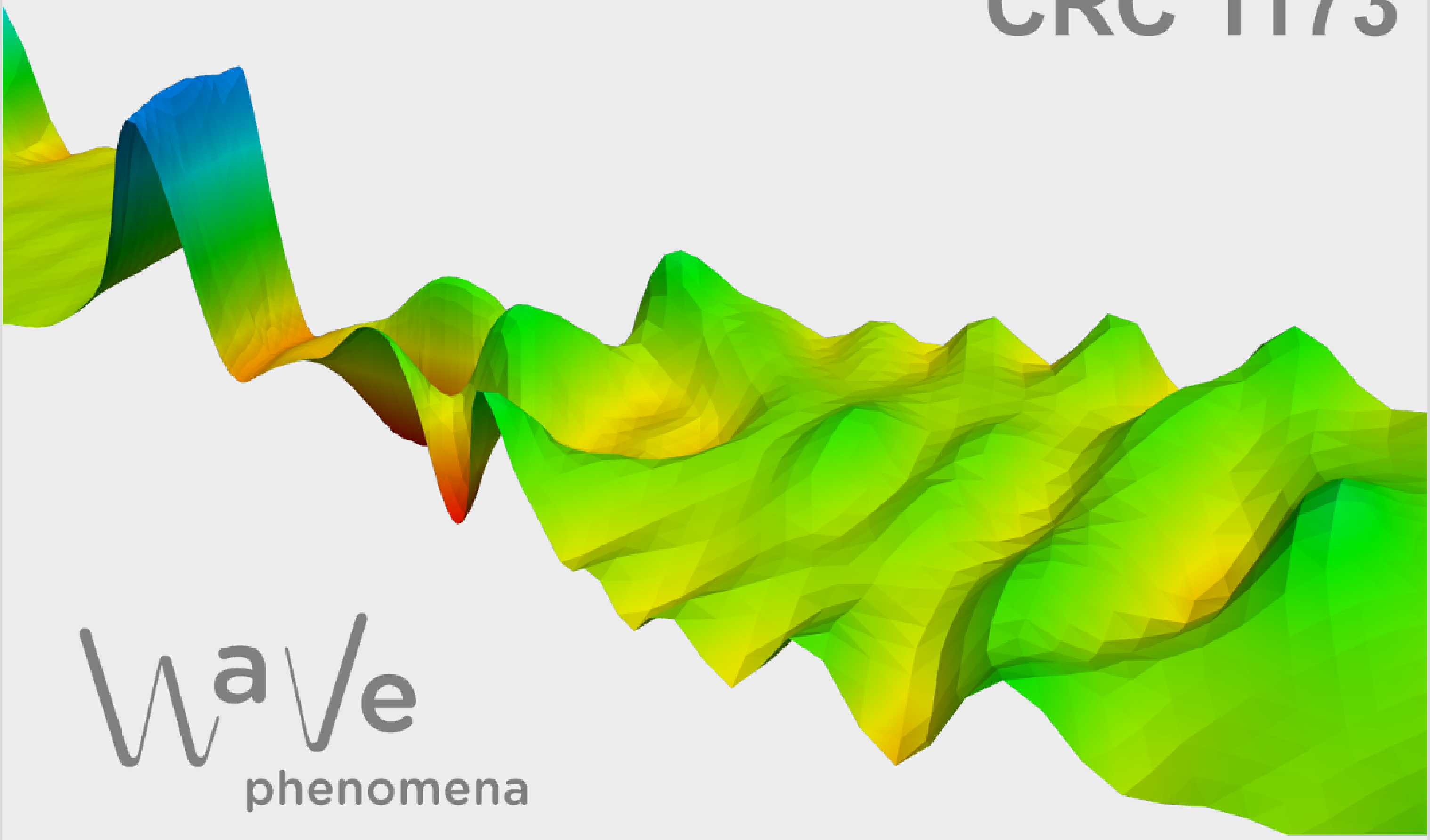
Tangential cone condition for the full waveform forward operator in the viscoelastic regime: the non-local case

Matthias Eller, Roland Griesmaier, Andreas Rieder

CRC Preprint 2023/8, February 2023



Participating universities

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

TU WIEN

Funded by

DFG

ISSN 2365-662X

TANGENTIAL CONE CONDITION FOR THE FULL WAVEFORM FORWARD OPERATOR IN THE VISCOELASTIC REGIME: THE NON-LOCAL CASE

MATTHIAS ELLER, ROLAND GRIESMAIER, AND ANDREAS RIEDER

Abstract. We discuss mapping properties of the parameter-to-state map of full waveform inversion and generalize the results of [M. Eller and A. Rieder, *Inverse Problems* 37 (2021) 085011] from the acoustic to the viscoelastic wave equation. In particular we establish injectivity of the Fréchet derivative of the parameter-to-state map for a semi-discrete seismic inverse problem in the viscoelastic regime. Here, the finite dimensional parameter space is restricted to functions having global support in the propagation medium (the non-local case) and that are locally linearly independent. As a consequence we deduce local conditional wellposedness of this nonlinear inverse problem. Furthermore, we show that the tangential cone condition holds, which is an essential prerequisite in the convergence analysis of a variety of inversion algorithms for nonlinear illposed problems.

1. Introduction

Time-domain full wave form inversion (FWI) aims to determine material parameters (such as mass density, shear and pressure wave speeds) of the earth's subsurface from reflection measurements of seismic wave fields (seismograms), using the full information content of seismic recordings, see, e.g., [22, 44]. In this work we discuss a theoretical aspect of FWI in the viscoelastic regime. Unlike purely elastic materials, viscoelastic materials are endowed with a memory in the sense that the state of stress at a certain instant of time depends on all deformations undergone by the material in previous times [15]. We consider the viscoelastic wave equation [43] in the velocity stress formulation based on the generalized standard linear solid rheology as described in [6, 40], see also [22].

Mathematically speaking, FWI is a nonlinear illposed parameter identification problem for the viscoelastic wave equation with partial measurements of viscoelastic waves. Usually these waves are initiated by controlled explosions, and the inverse problem is typically solved using Newton-like iterative regularization schemes, see, e.g., [9, 21, 37, 45]. The mathematical analysis of such schemes (see, e.g., [23, 29, 39]) relies crucially on a structural assumption on the nonlinear forward map known as the tangential cone condition (TCC, sometimes also referred to as the η-condition), which was introduced in [41]. A nonlinear operator $F\colon \mathsf{D}(F) \subset V \to W$ between normed spaces V and W satisfies the TCC at $x^+ \in \operatorname{int}(\mathsf{D}(F))$ if there are an $\eta \in (0,1)$ and an open ball $B_r(x^+) \subset \mathsf{D}(F)$ such

Date: February 7, 2023.

2010 *Mathematics Subject Classification.* 35R30, 35L05, 65J20.

Key words and phrases. Nonlinear illposed problem, tangential cone condition, Lipschitz stability, full waveform seismic inversion, viscoelastic wave equation.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 258734477 - SFB 1173.

that

$$\|F(v) - F(w) - F'(w)[v - w]\|_W \le \eta \, \|F(v) - F(w)\|_W \quad \text{for all } v, w \in B_r(x^+).$$

Here, $F' \colon \mathsf{D}(F) \subset V \to \mathcal{L}(V, W)$ denotes the Fréchet derivative of F. We refer to the monographs [20, 29, 42] and the recent publications [27, 28, 38] as evidence for the importance of TCC in the regularization theory of nonlinear illposed problems.

For infinite dimensional nonlinear illposed problems the TCC is often difficult to prove, and it has actually been established for only a few academic examples so far, see, e.g., [24, 29, 30]. In a semi-discrete setting for the inverse problem of electrical impedance tomography a TCC has been derived in [36]. In [19] the authors have shown that the TCC holds at x^+ if V is finite-dimensional (the semi-discrete situation) and the Fréchet derivative $F'(x^+)$ has a trivial null space. Therewith the TCC has been verified for full waveform inversion in the acoustic regime. In this work we use the abstract result from [19] to validate the TCC for the FWI forward operator in the viscoelastic regime provided the parameters (mass density, shear and pressure wave velocities with corresponding scaling factors) are restricted to a suitable finite-dimensional space, which is spanned by functions being analytic in the propagation medium (the non-local case).

Injectivity of the $F'(x^+)$ is sufficient not only for the TCC in the semi-discrete setting as described above but also for Lipschitz stability of the semi-discrete inverse problem. We note that independent of this observation Lipschitz estimates and associated conditional wellposedness for various semi-discrete inverse problems have recently attracted increasing interest, see, e.g., [1, 2, 3, 4, 5, 25]. For related works concerning the convergence analysis of regularization schemes for nonlinear inverse problems using conditional stability estimates we refer, e.g., to [13, 16, 31].

The outline of this work is as follows. We begin our presentation in the next section by introducing the viscoelastic wave equation as a first order hyperbolic system along with some statements on its wellposedness. Then, we formulate the forward operator Φ of the semi-discrete version of FWI in the viscoelastic regime. An important property is the Lipschitz continuity of the Fréchet derivative of this forward operator which we state in Theorem 2.3. The rather technical proof is moved to an appendix. In preparation for our main result in Theorem 4.3 we provide a control result for the viscoelastic wave equation in Section 3: given two open subsets Σ and Ω of the propagation medium, we can find a source in Σ such that the resulting velocity field at a sufficiently large time has non-trivial divergence and non-trivial deviator in Ω (see Theorem 3.4). The proof relies on a global Holmgren-John theorem for the homogeneous viscoelastic wave equation across non-characteristic surfaces from [18] (see also [12] for the corresponding local result). As a consequence of this controllability, the Fréchet derivative of Φ must be one-to-one at each inner point of the propagation medium (see Theorem 4.2). An application of Lemma C.1 of [19] finally yields the TCC for Φ and the Lipschitz-stability of the inverse problem. We conclude our work with a discussion of possible future research.

2. The setting

In two subsections we introduce the mathematical background of the considered forward and related inverse problem.

2.1. **The forward model.**

2.1.1. *The viscoelastic wave equation.* Wave propagation in realistic media can be modeled by a viscoelastic wave equation which accounts for dispersion and attenuation [22, Chap. 5]. Here, we derive the formulation introduced in [46] following the presentation from [9] and [14]: Let $D \subset \mathbb{R}^3$ be a bounded C^2 domain. The system of first-order wave equations for viscoelastic media describes the evolution of the particle velocity field $\boldsymbol{v}\colon [0,\infty) \times D \to \mathbb{R}^3$ and the stress tensor $\boldsymbol{\sigma}\colon [0,\infty) \times D \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$. It consists of the balance of momentum

$$\varrho(x)\, \partial_t \boldsymbol{v}(t,x) = \operatorname{div} \boldsymbol{\sigma}(t,x) + \boldsymbol{f}(t,x), \quad (t,x) \in [0,\infty) \times D,$$

and the retarded material law

$$\partial_t \boldsymbol{\sigma}(t,x) = \boldsymbol{C}(0)\varepsilon(\boldsymbol{v}(t,x)) + \int_0^t \partial_t \boldsymbol{C}(t-s)\varepsilon(\boldsymbol{v}(s,x))\,\mathrm{d}s + \boldsymbol{g}(t,x), \quad (t,x) \in [0,\infty) \times D,$$

where[1] $\boldsymbol{C}\colon [0,\infty) \to \mathcal{L}(\mathbb{R}^{3\times 3}_{\mathrm{sym}}, \mathbb{R}^{3\times 3}_{\mathrm{sym}})$ is the time-dependent Hooke tensor, $\boldsymbol{f}\colon [0,\infty) \times D \to \mathbb{R}^3$ is the volume force density, and $\boldsymbol{g}\colon [0,\infty) \times D \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$ is the time derivative of the external moment density. Further, $\varrho\colon D \to \mathbb{R}$ is the mass density and

$$\boldsymbol{\varepsilon}(\boldsymbol{v}) := \frac{1}{2}\big[(\nabla_x \boldsymbol{v})^\top + \nabla_x \boldsymbol{v}\big]$$

is the linearized strain rate.

In the *generalized standard linear solid* rheology using $L \in \mathbb{N}$ damping terms one defines

$$\boldsymbol{C}(t) := C(\mu,\pi) + \sum_{l=1}^{L} \exp\Big(-\frac{t}{\tau_{\boldsymbol{\sigma},l}}\Big) C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big) \tag{2.1}$$

with relaxation times $\tau_{\boldsymbol{\sigma},l} > 0$, $l = 1,\dots,L$. The functions $\tau_{\mathrm{P}}, \tau_{\mathrm{S}}\colon D \to \mathbb{R}$ are scaling factors for the unrelaxed bulk modulus $\pi\colon D \to \mathbb{R}$ and shear modulus $\mu\colon D \to \mathbb{R}$, respectively. The linear map $C(m,p)\colon \mathbb{R}^{3\times 3}_{\mathrm{sym}} \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$, $m,p \in \mathbb{R}$, models a Hooke element

$$C(m,p)\boldsymbol{M} := 2m\boldsymbol{M} + (p-2m)\operatorname{trace}(\boldsymbol{M})\boldsymbol{I}, \tag{2.2}$$

where $\boldsymbol{I} \in \mathbb{R}^{3\times 3}$ is the identity matrix and $\operatorname{trace}(\boldsymbol{M})$ denotes the trace of $\boldsymbol{M} \in \mathbb{R}^{3\times 3}_{\mathrm{sym}}$.

Introducing damping tensors $\boldsymbol{\sigma}_l\colon [0,\infty) \times D \to \mathbb{R}^{3\times 3}_{\mathrm{sym}}$,

$$\boldsymbol{\sigma}_l(t) := \int_0^t \exp\Big(\frac{s-t}{\tau_{\boldsymbol{\sigma},l}}\Big) C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big)\varepsilon(\boldsymbol{v}(s))\,\mathrm{d}s, \qquad l = 1,\dots,L,$$

and the corresponding stress decomposition $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + \sum_{l=1}^{L} \boldsymbol{\sigma}_l$, yields the first order system for viscoelastic waves

$$\varrho\, \partial_t \boldsymbol{v} = \operatorname{div}\Big(\sum_{l=0}^{L} \boldsymbol{\sigma}_l\Big) + \boldsymbol{f} \qquad \text{in } [0,\infty) \times D, \tag{2.3a}$$

$$\partial_t \boldsymbol{\sigma}_0 = C\big(\mu,\pi\big)\,\varepsilon(\boldsymbol{v}) + \boldsymbol{g} \qquad \text{in } [0,\infty) \times D, \tag{2.3b}$$

$$\tau_{\boldsymbol{\sigma},l}\, \partial_t \boldsymbol{\sigma}_l = \tau_{\boldsymbol{\sigma},l}\, C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big)\,\varepsilon(\boldsymbol{v}) - \boldsymbol{\sigma}_l, \qquad l = 1,\dots,L, \qquad \text{in } [0,\infty) \times D, \tag{2.3c}$$

[1] Throughout $\mathcal{L}(Y,Z)$ denotes the space of bounded linear operators between vector spaces Y and Z.

with initial conditions

$$\boldsymbol{v}(0) = \boldsymbol{v}_0 \quad \text{and} \quad \boldsymbol{\sigma}_l(0) = \boldsymbol{\sigma}_{l,0}, \quad l = 0, \dots, L. \tag{2.3d}$$

Wave propagation in viscoelastic media is frequency-dependent, and the relaxation times $\tau_{\boldsymbol{\sigma},l} > 0$ are used to model this dependency in a bounded frequency band with center frequency ω_0, see [7, 8]. Introducing the frequency-dependent phase velocities of P- and S-waves,

$$v_{\mathrm{P}}^2 := \frac{\pi}{\varrho}(1 + \tau_{\mathrm{P}}\alpha) \quad \text{and} \quad v_{\mathrm{S}}^2 := \frac{\mu}{\varrho}(1 + \tau_{\mathrm{S}}\alpha) \quad \text{with } \alpha := \sum_{l=1}^{L} \frac{\omega_0^2 \tau_{\boldsymbol{\sigma},l}^2}{1 + \omega_0^2 \tau_{\boldsymbol{\sigma},l}^2}, \tag{2.4}$$

full waveform inversion entails the identification of the five spatially dependent parameters $(\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}})$ from partial wave field measurements of $\boldsymbol{v}$ and $\boldsymbol{\sigma}$.

2.1.2. *Formulation as an abstract evolution equation.* Assuming that $\boldsymbol{g} = \mathbf{0}$, we rewrite (2.3) as an initial value problem for an abstract evolution equation. Our assumption on $\boldsymbol{g}$ is not a principal or strong restriction but eases the presentation somewhat and we can rely on the wellposedness results of [34]. For the general case we refer to [46].

Our presentation below closely follows [34]. Let

$$X := L^2(D, \mathbb{R}^3) \times L^2(D, \mathbb{R}^{3\times 3}_{\mathrm{sym}})^{1+L}$$

which is a Hilbert space when equipped with the inner product

$$\big\langle (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L), (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_l) \big\rangle_X := \int_D \Big(\boldsymbol{v} \cdot \mathbf{w} + \sum_{l=0}^{L} \boldsymbol{\sigma}_l : \boldsymbol{\psi}_l \Big)\, \mathrm{d}x$$

where the colon denotes the Frobenius inner product on $\mathbb{R}^{3\times 3}$.

For suitable[2] $w = (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L) \in X$ we define operators A, B, and Q mapping into X by

$$Aw := -\begin{pmatrix} \operatorname{div}\big(\sum_{l=0}^{L} \boldsymbol{\psi}_l\big) \\ \varepsilon(\boldsymbol{w}) \\ \vdots \\ \varepsilon(\boldsymbol{w}) \end{pmatrix}, \quad B^{-1}w := \begin{pmatrix} \frac{1}{\varrho}\boldsymbol{w} \\ C\big(\mu, \pi\big)\boldsymbol{\psi}_0 \\ C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big)\boldsymbol{\psi}_1 \\ \vdots \\ C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big)\boldsymbol{\psi}_L \end{pmatrix}, \quad Qw := \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \frac{1}{\tau_{\boldsymbol{\sigma},1}}\boldsymbol{\psi}_1 \\ \vdots \\ \frac{1}{\tau_{\boldsymbol{\sigma},L}}\boldsymbol{\psi}_L \end{pmatrix}. \tag{2.5}$$

Therewith, (2.3) can be reformulated as

$$B\partial_t u(t) + (A + BQ)u(t) = f(t), \quad t \in (0, \infty), \quad u(0) = u_0, \tag{2.6}$$

where

$$u(t) = \big(\boldsymbol{v}(t,\cdot), \boldsymbol{\sigma}_0(t,\cdot), \dots, \boldsymbol{\sigma}_L(t,\cdot)\big)^\top, \quad f(t) = \big(\boldsymbol{f}(t,\cdot), \mathbf{0}, \dots, \mathbf{0}\big)^\top,$$
$$\text{and} \quad u_0 = \big(\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0}\big)^\top.$$

In the remainder of this subsection we provide a domain $\mathsf{D}(A) \subset X$ for the differential operator A and we specify ranges for μ, π, τ_{S}, and τ_{P}, such that (2.6) is wellposed.

We define

$$\mathsf{D}(A) := H_0^1(D, \mathbb{R}^3) \times H(\operatorname{div})^{1+L}$$

[2]A rigorous description of the domains of these operators will be given below.

with
$$H(\mathrm{div}) = \big\{\boldsymbol{\sigma} \in L^2\big(D, \mathbb{R}^{3\times 3}_{\mathrm{sym}}\big) : \mathrm{div}\, \boldsymbol{\sigma}_{*,j} \in L^2(D),\ j = 1,2,3\big\}.$$
Then, the operator $A\colon \mathsf{D}(A) \subset X \to X$ is maximal monotone (see, e.g., [34, Lmm. 4.1]).

Now we consider the operator B. To this end we regard C of (2.2) as a mapping[3]
$$C\colon \mathsf{D}(C) \to \mathrm{Aut}(\mathbb{R}^{3\times 3}_{\mathrm{sym}}) \quad \text{where } \mathsf{D}(C) := \big\{(m,p) \in \mathbb{R}^2 : \underline{\mathrm{m}} \le m \le \overline{\mathrm{m}},\ \underline{\mathrm{p}} \le p \le \overline{\mathrm{p}}\big\}$$
with constants $0 < \underline{\mathrm{m}} < \overline{\mathrm{m}}$ and $0 < \underline{\mathrm{p}} < \overline{\mathrm{p}}$ such that $3\underline{\mathrm{p}} > 4\overline{\mathrm{m}}$. For $(m,p) \in \mathsf{D}(C)$,
$$C(m,p)^{-1} = C\left(\frac{1}{4m}, \frac{p-m}{m(3p-4m)}\right). \tag{2.7}$$
Moreover, $(C(m,p)\boldsymbol{M}) : \boldsymbol{N} = \boldsymbol{M} : (C(m,p)\boldsymbol{N})$ and
$$\min\{2\underline{\mathrm{m}}, 3\underline{\mathrm{p}} - 4\overline{\mathrm{m}}\}\, \boldsymbol{M} : \boldsymbol{M} \le C(m,p)\boldsymbol{M} : \boldsymbol{M} \le \max\{2\overline{\mathrm{m}}, 3\overline{\mathrm{p}} - 4\underline{\mathrm{m}}\}\, \boldsymbol{M} : \boldsymbol{M}, \tag{2.8}$$
see, e.g., [46, Lmm. 50].

As explained above, in FWI the five parameters $(\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}})$ are to be recovered and B depends on them via the relations
$$\pi = \frac{\varrho\, v_{\mathrm{P}}^2}{1 + \tau_{\mathrm{P}}\alpha} \quad \text{and} \quad \mu = \frac{\varrho\, v_{\mathrm{S}}^2}{1 + \tau_{\mathrm{S}}\alpha}, \tag{2.9}$$
see (2.4). Throughout we restrict these five parameters to the physically meaningful set
$$\begin{aligned}\mathcal{P} := \big\{\boldsymbol{p} = (p_1, \dots, p_5)^\top \in L^\infty(D)^5 : \varrho_{\min} < p_1(\cdot) < \varrho_{\max},\ v_{\mathrm{S},\min} < p_2(\cdot) < v_{\mathrm{S},\max},\\ \tau_{\mathrm{S},\min} < p_3(\cdot) < \tau_{\mathrm{S},\max},\ v_{\mathrm{P},\min} < p_4(\cdot) < v_{\mathrm{P},\max},\ \tau_{\mathrm{P},\min} < p_5(\cdot) < \tau_{\mathrm{P},\max} \text{ a.e. in } D\big\}\end{aligned} \tag{2.10}$$
with suitable bounds $0 < \varrho_{\min} < \varrho_{\max} < \infty$, etc. In view of (2.9) we deduce
$$\mu_{\min} := \frac{\varrho_{\min}\, v_{\mathrm{S},\min}^2}{1 + \tau_{\mathrm{S},\max}\alpha} \quad \text{and} \quad \mu_{\max} := \frac{\varrho_{\max}\, v_{\mathrm{S},\max}^2}{1 + \tau_{\mathrm{S},\min}\alpha} \tag{2.11}$$
as lower and upper bound on μ, and we observe that $\mu_{\min} < \mu < \mu_{\max}$ for all $(\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}}) \in \mathcal{P}$. The bounds $\pi_{\min}$ and $\pi_{\max}$ for π are defined accordingly by replacing S by P in (2.11).

Next we determine $\underline{\mathrm{p}}$, $\overline{\mathrm{p}}$, $\underline{\mathrm{m}}$, and $\overline{\mathrm{m}}$ such that $(\mu, \pi), (\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi) \in \mathsf{D}(C)$ whenever $(\varrho, v_{\mathrm{P}}, v_{\mathrm{S}}, \tau_{\mathrm{P}}, \tau_{\mathrm{S}}) \in \mathcal{P}$. In fact, this can be achieved by setting
$$\underline{\mathrm{p}} := \pi_{\min} \min\{1, \tau_{\mathrm{P},\min}\} \quad \text{and} \quad \overline{\mathrm{p}} := \pi_{\max} \max\{1, \tau_{\mathrm{P},\max}\}$$
with $\underline{\mathrm{m}}$ and $\overline{\mathrm{m}}$ defined correspondingly. The restriction $3\underline{\mathrm{p}} > 4\overline{\mathrm{m}}$ is equivalent to
$$\frac{4}{3}\, \frac{\varrho_{\max}}{\varrho_{\min}}\, \frac{1 + \tau_{\mathrm{P},\max}\alpha}{1 + \tau_{\mathrm{S},\min}\alpha}\, \frac{\max\{1, \tau_{\mathrm{S},\max}\}}{\min\{1, \tau_{\mathrm{P},\min}\}} < \frac{v_{\mathrm{P},\min}^2}{v_{\mathrm{S},\max}^2}. \tag{2.12}$$
We note that (2.12) implies that the family of selfadjoint operators $\{B = B(\boldsymbol{p}) : \boldsymbol{p} \in \mathcal{P}\} \subset \mathcal{L}(X)$ is uniformly positive definite and uniformly bounded.

Remark 2.1. *As there is no a priori physical relation/restriction between the scaling factors τ_{S} and τ_{P} it might happen locally that the shear wave travels faster than the pressure wave, see (2.4). In many applications in seismic and seismology it is often assumed that $\tau_{\mathrm{S}} \approx \tau_{\mathrm{P}}$.*[4] *In this case, relation (2.12) mathematically expresses the common observation that pressure waves propagate faster than shear waves.*

[3]We denote by $\mathrm{Aut}(\mathbb{R}^{3\times 3}_{\mathrm{sym}})$ the space of isomorphisms from $\mathbb{R}^{3\times 3}_{\mathrm{sym}}$ onto itself.

[4]Personal communication by Thomas Bohlen (Karlsruhe Institute of Technology, Geophysical Institute).

Finally, given the assumptions and restrictions from above, we can state the following wellposedness results: If $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top \in \mathsf{D}(A)$ and $\boldsymbol{f} \in W^{1,1}\big([0,\infty), L^2(D,\mathbb{R}^3)\big)$, then (2.6) (or equivalently (2.3) with $\boldsymbol{g} = \mathbf{0}$) admits a unique classical solution $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top \in \mathcal{C}([0,\infty), \mathsf{D}(A)) \cap \mathcal{C}^1([0,\infty), X)$ (see, e.g., [34]). On the other hand, if $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top \in X$ and $\boldsymbol{f} \in L^1_{\mathrm{loc}}([0,\infty), L^2(D,\mathbb{R}^3))$, then (2.6) (or equivalently (2.3) with $\boldsymbol{g} = \mathbf{0}$) admits a unique mild solution $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top \in \mathcal{C}([0,\infty), X)$, which satisfies

$$\int_0^t u(s)\mathrm{d}s \in \mathsf{D}(A), \qquad t \in [0,\infty), \tag{2.13a}$$

and

$$Bu(t) = Bu_0 - (A + BQ)\int_0^t u(s)\mathrm{d}s + \int_0^t f(s)\mathrm{d}s, \qquad t \in [0,\infty), \tag{2.13b}$$

(see, e.g., [26, Thm 2.15]).

Remark 2.2. *Below in Section 3 we will need to extend the (mild) solution of the homogeneous version of (2.6) to negative times. This extension is possible because A is skew-symmetric, i.e., $-A = A^*$, and A^* is maximal monotone (see* [33, Rem. 3.2]*). The homogeneous version of (2.6) for negative time reads*

$$-B\partial_t u(-t) + (A^* - BQ)u(-t) = 0, \quad t \in (0,\infty), \quad u(0) = u_0. \tag{2.14}$$

The same arguments as in [34, Sect. 3] *can be used to show that (2.14) admits a unique (mild) solution, which satisfies*

$$Bu(-t) = Bu_0 - (A^* - BQ)\int_0^t u(-s)\mathrm{d}s = Bu_0 + (A + BQ)\int_{-t}^0 u(s)\mathrm{d}s, \qquad t \in [0,\infty),$$

(see, e.g., [26, Thm 2.15]*). Therewith, the mild solution u, which satisfies (2.13) with $f = 0$, can be extended to negative times.*

2.2. **The (semi-discrete) inverse problem.** Let

$$V := \mathrm{span}\{\varphi_j \text{ analytic in } \overline{D} \;:\; j = 1, \dots, M\} \subset L^\infty(\overline{D}).$$

Note that the functions $\{\varphi_j : j = 1, \dots, M\}$ are locally linearly independent over D. This means that any linear combination that vanishes on a subset $O \subset D$ with positive Lebesgue measure must be trivial:

$$\mathrm{meas}(O) > 0 \text{ and } \sum_{j=1}^M a_j \varphi_j|_O = 0 \quad \Longrightarrow \quad a_j = 0,\ j = 1, \dots, M. \tag{2.15}$$

We write $\|\cdot\|_V := \|\cdot\|_{L^\infty(\overline{D})}$. Specific examples for spaces V with the required properties are polynomial spaces and spaces spanned by certain classes of radial basis functions (see [19, Sec. 3]).

In a seismic experiment, sources are fired at time zero in a non-empty open subset $\Sigma \subset D$ and the resulting wave fields are measured in a different non-empty open subset $\Omega \subset D$ until the observation time $T > 0$ has been reached. Accordingly, the measurements are in $\mathcal{C}([0,T], X_\Omega)$ where $X_\Omega := L^2(\Omega, \mathbb{R}^3) \times L^2(\Omega, \mathbb{R}^{3\times 3}_{\mathrm{sym}})$. For technical reasons, which will become clear in the proof of Theorem 2.3 below, we confine the prescribed sources to

$$W_0^{2,1}(\Sigma) := \big\{\boldsymbol{f} \in W^{2,1}\big([0,T], L^2(\Sigma,\mathbb{R}^3)\big) \;:\; \boldsymbol{f}(0) = \boldsymbol{f}'(0) = \mathbf{0}\big\}. \tag{2.16}$$

To formulate the corresponding semi-discrete inverse problem we set $V_+^5 := V^5 \cap \mathcal{P}$ and define the FWI forward operator (parameter-to-source-to-state map) by

$$\Phi\colon V_+^5 \subset V^5 \to \underbrace{\mathcal{L}\big(W_0^{2,1}(\Sigma), \mathcal{C}([0,T], X_\Omega)\big)}_{=:\, \mathcal{W}}, \quad (\varrho, v_\mathrm{S}, \tau_\mathrm{S}, v_\mathrm{P}, \tau_\mathrm{P}) \mapsto \big(\boldsymbol{f} \mapsto \Psi(u)\big), \tag{2.17}$$

where $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$ is the classical solution of (2.6) (or equivalently (2.3) with $\boldsymbol{g} = \mathbf{0}$) with initial values $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top$. Here,

$$\Psi\colon \mathcal{C}\big([0,T], X\big) \to \mathcal{C}\big([0,T], X_\Omega\big), \quad \Psi(\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L) := \Big(\boldsymbol{w}|_\Omega, \sum_{l=0}^{L} \boldsymbol{\psi}_l|_\Omega\Big)^\top,$$

models the measurement process. Note that $\sum_{l=0}^{L} \boldsymbol{\sigma}_l|_\Omega$, which is the second component of $\Psi(u)$, accounts for the stress related to $u|_\Omega$.

Now the semi-discrete seismic inverse problem (time-domain full waveform inversion) in the viscoelastic regime reads:

$$\text{Reconstruct } \boldsymbol{p} = (\varrho, v_\mathrm{S}, \tau_\mathrm{S}, v_\mathrm{P}, \tau_\mathrm{P}) \in V_+^5 \text{ from a measured version of } \Phi(\boldsymbol{p}). \tag{2.18}$$

We will verify in Theorem 4.3 below that in contrast to its infinite-dimensional version, which is locally illposed (see [34, Thm. 4.3]), the semidiscrete inverse problem is in fact locally wellposed and Lipschitz stable.

It has been established in [34, Thm. 4.4] that Φ is Fréchet-differentiable at any $\boldsymbol{p} = (\varrho, v_\mathrm{S}, \tau_\mathrm{S}, v_\mathrm{P}, \tau_\mathrm{P}) \in V_+^5$ with derivative $\Phi'\colon V_+^5 \subset V^5 \to \mathcal{L}(V^5, \mathcal{W})$ given by

$$\Phi'(\boldsymbol{p})[\boldsymbol{h}]\boldsymbol{f} = \Psi(\overline{u}), \qquad \boldsymbol{h} = (h_1, \dots, h_5) \in V^5, \ \boldsymbol{f} \in W_0^{2,1}(\Sigma),$$

where $\overline{u} = (\overline{\boldsymbol{v}}, \overline{\boldsymbol{\sigma}}_0, \dots, \overline{\boldsymbol{\sigma}}_L)^\top \in \mathcal{C}\big([0,T], X\big)$ denotes the mild solution of

$$\varrho\, \partial_t \overline{\boldsymbol{v}} = \operatorname{div}\Big(\sum_{l=0}^{L} \overline{\boldsymbol{\sigma}}_l\Big) - h_1\, \partial_t \boldsymbol{v}, \tag{2.19a}$$

$$\partial_t \overline{\boldsymbol{\sigma}}_0 = C(\mu, \pi)\varepsilon(\overline{\boldsymbol{v}}) + \Big(\frac{h_1}{\varrho}\, C(\mu, \pi) + \varrho\, C(\widetilde{\mu}, \widetilde{\pi})\Big)\varepsilon(\boldsymbol{v}), \tag{2.19b}$$

$$\begin{aligned} \partial_t \overline{\boldsymbol{\sigma}}_l = {}& C(\tau_\mathrm{S}\mu, \tau_\mathrm{P}\pi)\varepsilon(\overline{\boldsymbol{v}}) \\ & - \frac{1}{\tau_{\boldsymbol{\sigma},l}}\, \overline{\boldsymbol{\sigma}}_l + \Big(\frac{h_1}{\varrho}\, C(\tau_\mathrm{S}\mu, \tau_\mathrm{P}\pi) + \varrho\, C(\widehat{\mu}, \widehat{\pi})\Big)\varepsilon(\boldsymbol{v}), \quad l = 1, \dots, L, \end{aligned} \tag{2.19c}$$

with initial values

$$\overline{u}_0 = (\overline{\boldsymbol{v}}_0, \overline{\boldsymbol{\sigma}}_{0,0}, \dots, \overline{\boldsymbol{\sigma}}_{L,0})^\top = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top. \tag{2.19d}$$

Here $\boldsymbol{v}$ is the first component of the classical solution $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$ of (2.6) (or equivalently (2.3) with $\boldsymbol{g} = \mathbf{0}$) with initial values $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top$. Further,

$$\widetilde{\mu} := \frac{2 v_\mathrm{S}}{1 + \tau_\mathrm{S}\alpha}\, h_2 - \frac{\alpha\, v_\mathrm{S}^2}{(1 + \tau_\mathrm{S}\alpha)^2}\, h_3, \qquad \widetilde{\pi} := \frac{2 v_\mathrm{P}}{1 + \tau_\mathrm{P}\alpha}\, h_4 - \frac{\alpha\, v_\mathrm{P}^2}{(1 + \tau_\mathrm{P}\alpha)^2}\, h_5, \tag{2.19e}$$

$$\widehat{\mu} := \frac{2\tau_\mathrm{S}\, v_\mathrm{S}}{1 + \tau_\mathrm{S}\alpha}\, h_2 + \frac{v_\mathrm{S}^2}{(1 + \tau_\mathrm{S}\alpha)^2}\, h_3, \qquad \widehat{\pi} := \frac{2\tau_\mathrm{P}\, v_\mathrm{P}}{1 + \tau_\mathrm{P}\alpha}\, h_4 + \frac{v_\mathrm{P}^2}{(1 + \tau_\mathrm{P}\alpha)^2}\, h_5. \tag{2.19f}$$

The following property of Φ' will become important later. Its technical proof is given in Appendix A.

Theorem 2.3. *The map $\Phi'\colon V^5_+ \subset V^5 \to \mathcal{L}(V^5, \mathcal{W})$ is Lipschitz continuous, i.e.,*[5]

$$\|\Phi'(\boldsymbol{p}_1) - \Phi'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{W})} \lesssim \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{V^5}$$

for all $\boldsymbol{p}_1, \boldsymbol{p}_2 \in V^5_+$. The Lipschitz constant depends only on the observation time T and the bounds $\varrho_{\min}$, $\varrho_{\max}$, etc., that restrict the parameter range $\mathcal{P}$ in (2.10).

3. A control result for the viscoelastic wave equation

In Theorem 3.4 below we will establish the existence of a source $\boldsymbol{f} \in W_0^{2,1}(\Sigma)$, which plugged into (2.3a) initiates a velocity field $\boldsymbol{v}$ with non-trivial divergence and non-trivial deviator in Ω at time $T > 0$ provided that T is large enough.

We define the bounded linear operator

$$\mathcal{T}\colon L^2([0,T] \times \Sigma, \mathbb{R}^3) \to L^2_\varrho(\Omega, \mathbb{R}^3), \quad \boldsymbol{f} \mapsto \boldsymbol{v}(T,\cdot)|_\Omega, \tag{3.1}$$

where $\boldsymbol{v}$ is the first component of $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$, the classical solution of (2.6) (or equivalently (2.3) with $\boldsymbol{g} = \boldsymbol{0}$) with initial values $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top = (\boldsymbol{0}, \boldsymbol{0}, \dots, \boldsymbol{0})^\top$. Note that $\mathcal{T}$ is well defined as $\boldsymbol{v}$ is continuous in time, and bounded because $\|u\|_{\mathcal{C}([0,T],X)} \lesssim \|\boldsymbol{f}\|_{L^1([0,T],X)}$ (see [34, p. 2643]). The space $L^2_\varrho(\Omega, \mathbb{R}^3)$ is the same as $L^2(\Omega, \mathbb{R}^3)$ but with the ϱ-weighted inner product

$$\langle \boldsymbol{\psi}, \boldsymbol{w} \rangle_{L^2_\varrho(\Omega,\mathbb{R}^3)} := \langle \varrho\, \boldsymbol{\psi}, \boldsymbol{w} \rangle_{L^2(\Omega,\mathbb{R}^3)}.$$

Both spaces share the same topology by the restrictions on ϱ in (2.10).

Lemma 3.1. *The adjoint operator $\mathcal{T}^*$ of $\mathcal{T}$ from (3.1) is given by*

$$\mathcal{T}^*\colon L^2_\varrho(\Omega, \mathbb{R}^3) \to L^2([0,T] \times \Sigma, \mathbb{R}^3), \quad \boldsymbol{r} \mapsto \boldsymbol{w}|_{[0,T]\times\Sigma},$$

where $g = (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L)^\top \in \mathcal{C}([0,T], X)$ is the unique mild solution of the adjoint wave equation[6]

$$B\partial_t g(t) = (A + BQ)^* g(t), \quad t \in (0,T), \quad g(T) = (\chi_\Omega \boldsymbol{r}, \boldsymbol{0}, \dots, \boldsymbol{0})^\top. \tag{3.2}$$

Proof. Let $\boldsymbol{f} \in L^2([0,T] \times \Sigma, \mathbb{R}^3)$ and $\boldsymbol{r} \in L^2_\rho(\Omega, \mathbb{R}^3)$. To work with classical solutions we choose sequences $\{\boldsymbol{f}_k\}_k \subset W^{1,1}([0,T], L^2(D, \mathbb{R}^3))$ and $\{\boldsymbol{r}_k\}_k \subset H^1_0(D, \mathbb{R}^3)$ with $\boldsymbol{f}_k \to \chi_\Sigma \boldsymbol{f}$ in $L^2([0,T] \times D, \mathbb{R}^3)$ and $\boldsymbol{r}_k \to \chi_\Omega \boldsymbol{r}$ in $L^2_\varrho(D, \mathbb{R}^3)$. Furthermore, let $u_k = (\boldsymbol{v}_k, \boldsymbol{\sigma}_{0,k}, \dots, \boldsymbol{\sigma}_{L,k})^\top$ and $g_k = (\boldsymbol{w}_k, \boldsymbol{\psi}_{0,k}, \dots, \boldsymbol{\psi}_{L,k})^\top$ be the classical solutions of (2.6) with initial values $u_k(0) = (\boldsymbol{0}, \boldsymbol{0}, \dots, \boldsymbol{0})^\top$ and of (3.2), respectively, when replacing f by $f_k := (\boldsymbol{f}_k, \boldsymbol{0}, \dots, \boldsymbol{0})^\top$ and $\boldsymbol{r}$ by $\boldsymbol{r}_k$. We note that $\boldsymbol{w}_k \to \boldsymbol{w}$ in $L^2([0,T] \times D, \mathbb{R}^3)$ and $\boldsymbol{v}_k(T,\cdot) \to \boldsymbol{v}(T,\cdot)$ in $L^2(D, \mathbb{R}^3)$ (see [34, pp. 2643]).

Integration by parts yields

$$\begin{aligned}
\langle \boldsymbol{v}_k(T,\cdot), \boldsymbol{r}_k \rangle_{L^2_\varrho(\Omega,\mathbb{R}^3)} &= \langle \varrho\, \boldsymbol{v}_k(T,\cdot), \boldsymbol{w}_k(T,\cdot) \rangle_{L^2(\Omega,\mathbb{R}^3)} - \langle \varrho\, \boldsymbol{v}_k(0,\cdot), \boldsymbol{w}_k(0,\cdot) \rangle_{L^2(\Omega,\mathbb{R}^3)} \\
&= \langle B u_k(T), g_k(T) \rangle_X - \langle u_k(0), B g_k(0) \rangle_X \\
&= \int_0^T \langle B\partial_t u_k(t), g_k(t) \rangle_X \mathrm{d}t + \int_0^T \langle u_k(t), B\partial_t g_k(t) \rangle_X \mathrm{d}t \\
&= \int_0^T \langle -(A + BQ) u_k(t) + f_k(t), g_k(t) \rangle_X \mathrm{d}t
\end{aligned}$$

[5] The notation $E_1 \lesssim E_2$ indicates the existence of a generic constant $c > 0$ such that $E_1 \le c\, E_2$.

[6] Since $(A + BQ)^* = -A + BQ$, the same arguments as in [34, Sect. 3] can be used to show that (3.2) admits a unique mild solution.

$$+\int_0^T \langle u_k(t), (A+BQ)^* g_k(t)\rangle_X \mathrm{d}t$$

$$= \int_0^T \langle f_k(t), g_k(t)\rangle_X \mathrm{d}t = \langle \boldsymbol{f}_k, \boldsymbol{w}_k\rangle_{L^2([0,T]\times D,\mathbb{R}^3)}.$$

Passing to the limit as $k \to \infty$ verifies the assertion. □

Theorem 3.2. *Suppose that $\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}} \in \mathcal{P}$ are analytic in $\overline{D}$. Then the operator $\mathcal{T}$ defined in (3.1) has a dense range, provided that*

$$T/2 > \operatorname{dist}(x,\Sigma) := \inf_{y\in\Sigma} \operatorname{dist}(x,y) \qquad \textit{for all } x \in D. \tag{3.3}$$

Here, dist *denotes the Riemannian distance function in D, which is defined by*

$$\operatorname{dist}(x,y) := \inf_{\gamma} \int_a^b \frac{|\dot\gamma(t)|}{c(\gamma(t))}\,\mathrm{d}t, \qquad x,y \in D, \tag{3.4}$$

with

$$c^2(x) := \min\left\{ v_{\mathrm{S}}^2(x)\,\frac{1+L\tau_{\mathrm{S}}(x)}{1+\alpha\tau_{\mathrm{S}}(x)}, v_{\mathrm{P}}^2(x)\,\frac{1+L\tau_{\mathrm{P}}(x)}{1+\alpha\tau_{\mathrm{P}}(x)} \right\},$$

where the infimum ist taken over all $\mathcal{C}^1$-curves $\gamma\colon [a,b] \to D$ connecting x and y.

Before we will prove the theorem we discuss the physical meaning of (3.3). In view of (2.4) and assuming $\alpha \approx L$ (note that $\alpha \le L$), c is about the velocity of the slowest wave type. The condition says that the allotted observation time T has to be large enough such that even the slowest waves initiated in Σ can propagate back and forth through D within the measurement period. In a typical rock formation we expect c to be close to the shear wave velocity, $c \approx v_{\mathrm{S}}$, but see Remark 2.1.

Proof of Theorem 3.2. From functional analysis we know that $\mathcal{T}$ has a dense range if and only if $\mathcal{T}^*$ is injective.

Now, assume that $\mathcal{T}^*\boldsymbol{r} = \boldsymbol{0}$. Let $g = (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L)^\top \in \mathcal{C}([0,T], X)$ be the mild solution of the adjoint wave equation (3.2), i.e., $B\partial_t g = (A+BQ)^* g = (-A+BQ)g$ subject to $g(T) = (\chi_\Omega \boldsymbol{r}, \boldsymbol{0}, \dots, \boldsymbol{0})^\top$. By the same argument as used in Remark 2.2, g can be extended to a solution of (3.2) that exists for all times $t \in \mathbb{R}$. Denoting this extension again by $g = (\boldsymbol{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_L)^\top \in \mathcal{C}(\mathbb{R}, X)$ we have that

$$\varrho\,\partial_t \boldsymbol{w} = \operatorname{div}\Big(\sum_{l=0}^{L} \boldsymbol{\psi}_l\Big) \qquad \text{in } \mathbb{R}\times D, \tag{3.5a}$$

$$\partial_t \boldsymbol{\psi}_0 = C\big(\mu, \pi\big)\,\varepsilon(\boldsymbol{w}) \qquad \text{in } \mathbb{R}\times D, \tag{3.5b}$$

$$\tau_{\boldsymbol{\sigma},l}\,\partial_t \boldsymbol{\psi}_l = \tau_{\boldsymbol{\sigma},l}\, C\big(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi\big)\,\varepsilon(\boldsymbol{w}) + \boldsymbol{\psi}_l, \qquad l = 1,\dots,L, \qquad \text{in } \mathbb{R}\times D, \tag{3.5c}$$

together with the initial conditions $\boldsymbol{w}(T,\cdot) = \boldsymbol{r}$ and $\boldsymbol{\psi}_l(T,\cdot) = \boldsymbol{0}$, $l = 0,\dots,L$, in D. This mild solution is the weak solution as well (see, e.g., [33, Cor. 2.5]). From Lemma 3.1 we find that our assumption $\mathcal{T}^*\boldsymbol{r} = \boldsymbol{0}$ means that $\boldsymbol{w} = 0$ everywhere in $[0,T]\times\Sigma$. Accordingly, (3.5b) together with the homogeneous initial condition implies then that $\boldsymbol{\psi}_0 = \boldsymbol{0}$ in $[0,T]\times\Sigma$. The L equations in (3.5c) can be viewed as linear first-order ordinary differential equations for $\boldsymbol{\psi}_l$. Taking into account the homogeneous initial data we find that

$$\boldsymbol{\psi}_l(t) = -\int_t^T \exp\left(\frac{t-s}{\tau_{\boldsymbol{\sigma},l}}\right) C(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(s))\,\mathrm{d}s, \quad l = 1,2,\dots,L, \quad t \in \mathbb{R}. \tag{3.6}$$

Hence, $\boldsymbol{w}|_{[0,T]\times\Sigma} = 0$ gives $\boldsymbol{\psi}_l = \mathbf{0}$ in $[0,T]\times\Sigma$ for all $l = 1,2,\dots,L$. Consequently we have shown that $g = (\mathbf{0},\dots,\mathbf{0})^\top$ in $[0,T]\times\Sigma$.

By taking the divergence of (3.5b), the time derivative of (3.5a), and using the explicit expressions (3.6) for the $\boldsymbol{\psi}_l$'s, we obtain a second-order system with memory term

$$\begin{aligned}\varrho\partial_t^2\boldsymbol{w}(t) &= \operatorname{div}\big(C(\mu,\pi)\varepsilon(\boldsymbol{w}(t))\big) + \sum_{l=1}^L \operatorname{div}\Bigg[C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(t)) \\ &\qquad - \frac{1}{\tau_{\boldsymbol{\sigma},l}}\int_t^T \exp\left(\frac{t-s}{\tau_{\boldsymbol{\sigma},l}}\right) C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(s))\,\mathrm{d}s\Bigg] \\ &= \operatorname{div}\big(C\left([1+L\tau_{\mathrm{S}}]\mu,[1+L\tau_{\mathrm{P}}]\pi\right)\varepsilon(\boldsymbol{w}(t))\big) \\ &\qquad - \operatorname{div}\int_t^T \left[\sum_{l=1}^L \frac{1}{\tau_{\boldsymbol{\sigma},l}}\exp\left(\frac{t-s}{\tau_{\boldsymbol{\sigma},l}}\right)\right] C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(s))\,\mathrm{d}s.\end{aligned} \tag{3.7}$$

Accordingly, $\widetilde{\boldsymbol{w}}(t) := \boldsymbol{w}(-t-T)$, $t\in\mathbb{R}$, satisfies a wave equation that is exactly of the form discussed in formulas (1) and (2) of [18] with relaxation tensor

$$H(t) = C(\mu,\pi) + \sum_{j=1}^L \exp\left(-\frac{t}{\tau_{\boldsymbol{\sigma},l}}\right) C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi),$$

using the notation employed in this reference. Indeed, we verify that

$$H(0) = C(\mu,\pi) + \sum_{l=1}^L C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi) = C\big([1+L\tau_S]\mu,[1+L\tau_P]\pi\big),$$

and that

$$-\partial_t H(t) = \sum_{l=1}^L \frac{1}{\tau_{\boldsymbol{\sigma},l}}\exp\left(-\frac{t}{\tau_{\boldsymbol{\sigma},l}}\right) C(\tau_{\mathrm{S}}\mu,\tau_{\mathrm{P}}\pi).$$

We would like to invoke a global uniqueness statement of [18, Thm 1.4] which, however, requires solutions to the system (3.7) of class H^2, i.e., $\boldsymbol{w}\in H^2([0,T]\times D,\mathbb{R}^3)$. By a regularization argument, we will demonstrate that the local uniqueness property holds also for mild solutions $\boldsymbol{w}\in\mathcal{C}(\mathbb{R},L^2(D,\mathbb{R}^3))$. To this end let χ be the indicator function of the interval $[-1/2,1/2]$ and define $\chi_n(\cdot) := n\chi(n\,\cdot)$ for $n\in\mathbb{N}$. Further, set $g_n := \chi_n\star\chi_n\star g$ where $\star$ denotes convolution in time. The regularizations $g_n = (\boldsymbol{w}_n,\boldsymbol{\psi}_{0,n},\dots,\boldsymbol{\psi}_{L,n})$ and their time derivatives are solutions to the system (3.5) in $\mathbb{R}\times D$ since the coefficients are time-independent. Additionally, $g_n(t)\in\mathsf{D}(A)$ for $t\in\mathbb{R}$ which follows recursively from

$$\frac{1}{n}\,\chi_n\star g(t) = \int_{t-1/(2n)}^{t+1/(2n)} g(s)\mathrm{d}s = \int_0^{t+1/(2n)} g(s)\mathrm{d}s - \int_0^{t-1/(2n)} g(s)\mathrm{d}s$$

since the latter two functions are in $\mathsf{D}(A)$ as g is a mild solution of (3.5) (see (2.13a) which extends to negative times according to Remark 2.2). In particular,

$$\boldsymbol{w}_n(T)\in H_0^1(D,\mathbb{R}^3) \quad\text{and}\quad \operatorname{div}\boldsymbol{\psi}_{l,n}(T)\in L^2(D,\mathbb{R}^3),\ l=0,\dots,L. \tag{3.8}$$

Moreover, $g_n\in\mathcal{C}^2(\mathbb{R},X)$ since $g\in\mathcal{C}(\mathbb{R},X)$. Accordingly, $\partial_t^2\boldsymbol{w}_n(T)\in L^2(D,\mathbb{R}^3)$. The relation

$$\frac{1}{n}\,\partial_t g_n(t) = \chi_n\star g(t+1/(2n)) - \chi_n\star g(t-1/(2n))$$

proves that $\partial_t g_n(t) \in \mathsf{D}(A)$ for all $t \in \mathbb{R}$. Thus, in view of (3.5a) for g_n,

$$\operatorname{div} \sum_{l=0}^{L} \boldsymbol{\psi}_{l,n}(T) \in H_0^1(D, \mathbb{R}^3). \tag{3.9}$$

Evaluating (3.7) for $\boldsymbol{w}_n$ in $t = T$ gives

$$\operatorname{div}\left(C\left([1 + L\tau_{\mathrm{S}}]\mu, [1 + L\tau_{\mathrm{P}}]\pi\right) \varepsilon(\boldsymbol{w}_n(T))\right) = \varrho \partial_t^2 \boldsymbol{w}_n(T) \quad \text{in } D.$$

Elliptic regularity shows that $\boldsymbol{w}_n(T) \in H^2(D, \mathbb{R}^3)$ (see [10, Thm. 6.3-6]). Recalling that g_n satisfies (3.5), we have just shown that the final state $g_n(T)$ is in

$$\mathsf{D}(A^2) = \Big\{ (\boldsymbol{z}, \boldsymbol{\phi}_0, \dots, \boldsymbol{\phi}_L)^\top \in H_0^1(D, \mathbb{R}^3) \times H(\operatorname{div})^{1+L} : \\ \operatorname{div} \varepsilon(\boldsymbol{z}) \in L^2(D, \mathbb{R}^3),\ \operatorname{div} \sum_{l=0}^{L} \boldsymbol{\phi}_l \in H_0^1(D, \mathbb{R}^3) \Big\}.$$

In fact, $g_n(T) \in \mathsf{D}(A^2)$ can be inferred from (3.8), (3.9), and $\operatorname{div} \varepsilon(\boldsymbol{w}_n(T)) \in L^2(D, \mathbb{R}^3)$ as $\boldsymbol{w}_n(T) \in H^2(D, \mathbb{R}^3)$.

Setting $\widetilde{g}_n(t) := g_n(T - t)$, $t \in [0, T]$, and $\widetilde{A} := B^{-1}A - Q$, (3.5) can be written as initial value problem $\partial_t \widetilde{g}_n = \widetilde{A}\widetilde{g}_n$, $\widetilde{g}_n(0) := g_n(T)$. Further, $\partial_t \widetilde{g}_n(t) = S(t) g_n(T)$ where $\{S(t)\}_{t \geq 0}$ is the semigroup generated by $\widetilde{A}$. Since $\mathsf{D}(\widetilde{A}) = \mathsf{D}(A)$, $S(t)\mathsf{D}(\widetilde{A}^2) \subset \mathsf{D}(\widetilde{A}^2)$, and $\partial_t^2 S(t) g_n(T) = S(t)\widetilde{A}^2 g_n(T)$, see, e.g., [26, Prop. 2.11], we conclude that $g_n \in \mathcal{C}^2([0, T], \mathsf{D}(A^2))$. In particular, $\boldsymbol{w}_n \in H^2([0, T] \times D, \mathbb{R}^3)$.

Since $\boldsymbol{w} \equiv \mathbf{0}$ in $[0, T] \times \Sigma$, we know that $\boldsymbol{w}_n = \chi_n \star \chi_n \star \boldsymbol{w} \equiv \mathbf{0}$ in $[1/n, T - 1/n] \times \Sigma$. By assumption, there exists a $\delta > 0$ such that $T - 4\delta > 2\operatorname{dist}(x, \Sigma)$ for all $x \in D$. We will now show that we can use a global uniqueness theorem for solutions to the homogeneous viscoelastic system (Theorem 1.4 in [18]) to infer that $\boldsymbol{w}_n \equiv \mathbf{0}$ in $(T/2 - \delta, T/2 + \delta) \times D$ provided $n > 1/\delta$. For that we have to show that our distance function is indeed the distance function used in [18].

We will now work with the second-order system (3.7). Using the assumptions on the elasticity tensor, the characteristic polynomial is

$$\begin{aligned} &p(x, \tau, \xi) \\ &= \det\left(\tau^2(x) - \frac{\mu(x)[1 + L\tau_S(x)]}{\varrho(x)} |\xi|^2 I_3 - \frac{\pi(x)[1 + L\tau_P(x)] - \mu(x)[1 + L\tau_S(x)]}{\varrho(x)} \xi\xi^\top\right) \\ &= \left(\tau^2 - \frac{\mu(x)[1 + L\tau_S(x)]}{\varrho(x)} |\xi|^2\right)^2 \left(\tau^2 - \frac{\pi(x)[1 + L\tau_P(x)]}{\varrho(x)} |\xi|^2\right), \end{aligned}$$

see [18, Def. 1.1]. Hence, we have that

$$a_1(x, \xi) = \min\left\{\frac{\mu(x)[1 + L\tau_S(x)]}{\varrho(x)}, \frac{\pi(x)[1 + L\tau_P(x)]}{\varrho(x)}\right\} |\xi|^2$$

where a_1 is defined in [18, p. 1531]. Note that this function is of the form $c^2(x)|\xi|^2$, and c^2 is continuous. Hence, the distance function is defined by the metric $|v|_x = |v|/c(x)$ (see [18, p. 1532]).

Having established that $\boldsymbol{w}_n \equiv \mathbf{0}$ in $(T/2 - \delta, T/2 + \delta) \times D$ for all $n > 1/\delta$, by the convergence of the regularizations, the same must be true for the limit function $\boldsymbol{w}$.

Finally, we will show that $\boldsymbol{w}(T) = \boldsymbol{0}$ in D. To this end we consider again the second-order formulation of (3.5) similar to (3.7), but now we work with inital values at time $t = T/2$. In this case (3.6) has to be replaced by

$$\begin{aligned}\boldsymbol{\psi}_l(t) &= \exp\left(\frac{t - T/2}{\tau_{\boldsymbol{\sigma},l}}\right)\boldsymbol{\psi}_l(T/2) \\ &\quad + \int_{T/2}^{t} \exp\left(\frac{t-s}{\tau_{\boldsymbol{\sigma},l}}\right) C(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(s))\,\mathrm{d}s, \quad l = 1, 2, \ldots, L, \quad t \in \mathbb{R},\end{aligned} \tag{3.10}$$

and consequently

$$\begin{aligned}\varrho\partial_t^2\boldsymbol{w}(t) &= \operatorname{div}\left(C\left([1 + L\tau_{\mathrm{S}}]\mu, [1 + L\tau_{\mathrm{P}}]\pi\right)\varepsilon(\boldsymbol{w}(t))\right) \\ &\quad + \operatorname{div}\int_{T/2}^{t}\left[\sum_{l=1}^{L}\frac{1}{\tau_{\boldsymbol{\sigma},l}}\exp\left(\frac{t-s}{\tau_{\boldsymbol{\sigma},l}}\right)\right] C(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\varepsilon(\boldsymbol{w}(s))\,\mathrm{d}s \\ &\quad + \sum_{l=1}^{L}\exp\left(\frac{t - T/2}{\tau_{\boldsymbol{\sigma},l}}\right)\operatorname{div}\boldsymbol{\psi}_l(T/2).\end{aligned} \tag{3.11}$$

Using now that $\boldsymbol{w} \equiv 0$ in $(T/2 - \delta, T/2 + \delta) \times D$, gives

$$\sum_{l=1}^{L}\exp\left(\frac{t - T/2}{\tau_{\boldsymbol{\sigma},l}}\right)\operatorname{div}\boldsymbol{\psi}_l(T/2) = 0 \qquad \text{in } (T/2 - \delta, T/2 + \delta) \times D,$$

and since this function is analytic in t, we infer that this function must vanish for all $t \in \mathbb{R}$. Hence, in view of (3.11) the function $\boldsymbol{w}$ satisfies an elastic wave equation with memory term with vanishing initial data at $t = T/2$. This problem has a unique solution [11] and thus $\boldsymbol{w} \equiv \boldsymbol{0}$ in $(T/2, T) \times D$. In particular, $\boldsymbol{r} = \boldsymbol{w}(T) = \boldsymbol{0}$ in D. $\square$

Remark 3.3. *Theorem* 1.4 *in* [18] *uses a different setting from ours. The emphasis is on the uniqueness of the lateral Cauchy problem. Here we propagate a zero set given in the interior which is somewhat simpler. Observe that the Cauchy problem in* [18] *is treated by extending the function by zero to a larger set which means that the uniqueness question for the lateral Cauchy problem is actually converted into a uniqueness problem in the interior.*

Below we will need the deviator of a vector field. An element $\boldsymbol{\delta} \in L^2(\Omega, \mathbb{R}^{3\times3}_{\mathrm{sym}})$ is called the (weak) deviator of $\boldsymbol{w} \in L^2(\Omega, \mathbb{R}^3)$ if

$$\int_\Omega \boldsymbol{\delta} : \boldsymbol{\phi}\,\mathrm{d}x = -\int_\Omega \boldsymbol{w}\cdot\left(\operatorname{div}\boldsymbol{\phi} - \frac{1}{3}\nabla\operatorname{trace}(\boldsymbol{\phi})\right)\mathrm{d}x \qquad \text{for all } \boldsymbol{\phi} \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^{3\times3}_{\mathrm{sym}}).$$

We write $\operatorname{dev}\boldsymbol{w} := \boldsymbol{\delta}$ and note that $\operatorname{dev}\boldsymbol{w} = \varepsilon(\boldsymbol{w}) - \frac{1}{3}(\operatorname{div}\boldsymbol{w})\boldsymbol{I}_3$ for $\boldsymbol{w} \in H^1(\Omega, \mathbb{R}^3)$, i.e., $\operatorname{dev}\boldsymbol{w}$ is the trace-free part of $\varepsilon(\boldsymbol{w})$.

Theorem 3.4. *Assume that the observation time T satisfies* (3.3) *for the distance function* (3.4) *where c is replaced by its lower bound $\underline{c}$ with*[7]

$$\underline{c}^2 = \min\left\{v_{\mathrm{S,min}}^2\frac{1 + L\tau_{\mathrm{S,min}}}{1 + \alpha\tau_{\mathrm{S,min}}}, v_{\mathrm{P,min}}^2\frac{1 + L\tau_{\mathrm{P,min}}}{1 + \alpha\tau_{\mathrm{P,min}}}\right\}.$$

Then, there exist finitely many sources $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m \in W_0^{2,1}(\Sigma)$ for which the following holds: For any $\boldsymbol{p} = (\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}}) \in V_+^5$, there is an $\boldsymbol{f} \in \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m\}$ such that the

[7]Here we use (2.4) and the fact that $\alpha < L$.

first component of the classical solution $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$ *of* (2.6) *with source* $\boldsymbol{f}$ *and initial values* $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top$ *has a non-trivial divergence and a non-trivial deviator in* Ω *at time* T*, i.e.,* $\operatorname{div} \boldsymbol{v}(T,\cdot)|_\Omega \not\equiv 0$ *and* $\operatorname{dev} \boldsymbol{v}(T,\cdot)|_\Omega \not\equiv \mathbf{0}$.

Proof. Let $\boldsymbol{u} \in L^2(\Omega, \mathbb{R}^3)$ have non-trivial divergence and non-trivial deviator. Such a $\boldsymbol{u}$ can be constructed by choosing $\boldsymbol{u} = \nabla\varphi$ for some $\varphi \in \mathcal{C}^\infty(\overline{\Omega})$ with $\Delta\varphi \not\equiv 0$ and $\varepsilon_{ij}(\nabla\varphi) = \partial_{x_i}\partial_{x_j}\varphi \not\equiv 0$ for one pair (i,j) with $1 \le i \ne j \le 3$. For instance we can choose $\varphi(x) = \exp(-|x|^2)$, $x \in \Omega$.

Since the spaces of (weak) divergence free and (weak) deviator free vector fields in Ω are proper closed subspaces of $L^2(\Omega, \mathbb{R}^3)$ there is a ball $B_\varepsilon(\boldsymbol{u}) \subset L^2(\Omega, \mathbb{R}^3)$ around $\boldsymbol{u}$ with radius $\varepsilon > 0$ containing only vector fields with non-trivial divergence and non-trivial deviator. By Theorem 3.2, there is, for any $\boldsymbol{p} \in V_+^5$, an $\boldsymbol{f}_{\boldsymbol{p}} \in L^2([0,T] \times \Sigma, \mathbb{R}^3)$ such that $\mathcal{T}\boldsymbol{f}_{\boldsymbol{p}} \in B_{\varepsilon/2}(\boldsymbol{u})$, i.e., $\operatorname{div}(\mathcal{T}\boldsymbol{f}_{\boldsymbol{p}}) \not\equiv 0$ and $\operatorname{dev}(\mathcal{T}\boldsymbol{f}_{\boldsymbol{p}}) \not\equiv \mathbf{0}$. Note that $\mathcal{T}\boldsymbol{f}_{\boldsymbol{p}} = \Pi(\Phi(\boldsymbol{p})\boldsymbol{f}_{\boldsymbol{p}})(T,\cdot)$ where $\Pi\colon L^2(\Omega,\mathbb{R}^3) \times L^2(\Omega, \mathbb{R}^{3\times3}_{\text{sym}}) \to L^2(\Omega,\mathbb{R}^3)$ picks the first component. We can even choose $\boldsymbol{f}_{\boldsymbol{p}} \in W_0^{2,1}(\Sigma)$ (see (2.16)) because this space is dense in $L^2([0,T] \times \Sigma, \mathbb{R}^3)$. Additionally, we can assume the sources to be normalized: $\|\boldsymbol{f}_{\boldsymbol{p}}\|_{W_0^{2,1}(\Sigma)} = 1$.

For any other $\boldsymbol{q} \in V_+^5$ we have

$$\begin{aligned}
&\|\Pi(\Phi(\boldsymbol{p})\boldsymbol{f}_{\boldsymbol{q}})(T,\cdot) - \boldsymbol{u}\|_{L^2(\Omega,\mathbb{R}^3)} \\
&\qquad\qquad \le \|\Pi(\Phi(\boldsymbol{q})\boldsymbol{f}_{\boldsymbol{q}})(T,\cdot) - \boldsymbol{u}\|_{L^2(\Omega,\mathbb{R}^3)} + \|(\Phi(\boldsymbol{p}) - \Phi(\boldsymbol{q}))\boldsymbol{f}_{\boldsymbol{q}}\|_{\mathcal{C}([0,T],X_\Omega)} \\
&\qquad\qquad \le \frac{\varepsilon}{2} + \|(\Phi(\boldsymbol{p}) - \Phi(\boldsymbol{q}))\boldsymbol{f}_{\boldsymbol{q}}\|_{\mathcal{C}([0,T],X_\Omega)}.
\end{aligned}$$

Further, by [33, Lmm. 3.3] (this reference deals with the special case $Q = 0$, but the result immediately extends to nonzero Q as considered in this work; see also [34, Thm. 3.7]),

$$\|(\Phi(\boldsymbol{p}) - \Phi(\boldsymbol{q}))\boldsymbol{f}_{\boldsymbol{q}}\|_{\mathcal{C}([0,T],X_\Omega)} \le C_\Phi \|\boldsymbol{p} - \boldsymbol{q}\|_{V^5} \|\boldsymbol{f}_{\boldsymbol{q}}\|_{W_0^{2,1}(\Sigma)}$$

where the involved constant C_Φ depends on T and the bounds determining $\mathcal{P}$, see (2.10). Combining these two estimates we conclude that

$$\|\Pi(\Phi(\boldsymbol{p})\boldsymbol{f}_{\boldsymbol{q}})(T,\cdot) - \boldsymbol{u}\|_{L^2(\Omega,\mathbb{R}^3)} < \varepsilon/2 + C_\Phi\|\boldsymbol{p} - \boldsymbol{q}\|_{V^5}.$$

Since V_+^5 is compact we can cover it with finitely many, say m, open balls with radii less than $\varepsilon/(2C_\Phi)$. Denote the centers of these balls by $\boldsymbol{q}_1, \dots, \boldsymbol{q}_m$. Accordingly, there is a $\boldsymbol{q}_k$ such that $\|\boldsymbol{p} - \boldsymbol{q}_k\|_{V^5} < \varepsilon/(2C_\Phi)$. Hence,

$$\|\Pi(\Phi(\boldsymbol{p})\boldsymbol{f}_{\boldsymbol{q}_k})(T,\cdot) - \boldsymbol{u}\|_{L^2(\Omega,\mathbb{R}^3)} < \varepsilon$$

and both, divergence and deviator of $\Pi\Phi(\boldsymbol{p})\boldsymbol{f}_{\boldsymbol{q}_k}(T,\cdot)$, are not identically zero in Ω. □

4. Local injectivity of the FWI forward operator yields TCC

Proposition 4.1. *Suppose that* $\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}} \in \mathcal{P}$ *are analytic in* $\overline{D}$*. Let* $\boldsymbol{h} \in V^5 \backslash \{\mathbf{0}\}$*, and let* $\Sigma, \Omega \subset D$ *be non-empty, open and disjoint. If the observation time* T *is as in Theorem* 3.4*, then there exists an* $\boldsymbol{f} \in W_0^{2,1}(\Sigma)$ *such that* $\Psi(\overline{u}) \ne (\mathbf{0}, \mathbf{0})^\top$ *where* $\overline{u} = (\overline{\boldsymbol{v}}, \overline{\boldsymbol{\sigma}}_0, \dots, \overline{\boldsymbol{\sigma}}_L)^\top$ *is the mild solution of* (2.19) *with* $\boldsymbol{v}$ *being the first component of the classical solution* $u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top$ *of* (2.6) *with source* $\boldsymbol{f}$ *and initial values* $u_0 = (\boldsymbol{v}_0, \boldsymbol{\sigma}_{0,0}, \dots, \boldsymbol{\sigma}_{L,0})^\top = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top$*. This* $\boldsymbol{f}$ *does not depend on* $\boldsymbol{h}$.

Proof. We argue by contradiction. Assume that there exists an $\boldsymbol{h} = (h_1, h_2, h_3, h_4, h_5)^\top \in V^5 \backslash \{\mathbf{0}\}$ such that for any $\boldsymbol{f} \in W_0^{2,1}(\Sigma)$ the corresponding mild solution $\overline{u}$ of (2.19), where $\boldsymbol{v}$ is the first component of the classical solution u of (2.6) with source $\boldsymbol{f}$ and initial values $u_0 = (\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^\top$, satisfies $\Psi(\overline{u}) = (\mathbf{0}, \mathbf{0})$. According to Theorem 3.4 there is an $\boldsymbol{f}$ such that $\boldsymbol{v}$ has a non-trivial divergence and a non-trivial deviator in Ω at time T, i.e., $\operatorname{div} \boldsymbol{v}(T, \cdot)|_\Omega \not\equiv 0$ and $\operatorname{dev} \boldsymbol{v}(T, \cdot)|_\Omega \not\equiv \mathbf{0}$. In particular $\boldsymbol{v}(T, \cdot)|_\Omega \not\equiv \mathbf{0}$.

Note that $\overline{u} = (\overline{\boldsymbol{v}}, \overline{\boldsymbol{\sigma}}_0, \dots, \overline{\boldsymbol{\sigma}}_L)^\top$ satisfies the differential equation (2.19) as a mild solution only in the integrated form, compare (2.13). Using $\overline{\boldsymbol{v}}|_\Omega = \mathbf{0}$ and $\sum_{l=0}^L \overline{\boldsymbol{\sigma}}_l|_\Omega = \mathbf{0}$, the first equation of the integrated version of (2.19) yields

$$h_1(x)\, \partial_t \boldsymbol{v}(t, x) = \mathbf{0}, \quad (t, x) \in [0, T] \times \Omega.$$

If $h_1 \neq 0$ then $h_1(\cdot) \neq 0$ in Ω by (2.15). Hence, the above displayed equation implies that $\partial_t \boldsymbol{v}|_{[0,T]\times\Omega} = \mathbf{0}$ almost everywhere. In view of $\boldsymbol{v}(0, \cdot) = \boldsymbol{v}_0 = \mathbf{0}$ we get that $\boldsymbol{v}(t, \cdot)|_\Omega = \mathbf{0}$ almost everywhere for any $t \in [0, T]$ contradicting $\boldsymbol{v}(T, \cdot)|_\Omega \not\equiv \mathbf{0}$. Therefore, $h_1 = 0$.

Next we consider the remaining equations of the integrated version of (2.19). They yield

$$\begin{aligned}
&\partial_t \overline{\boldsymbol{\sigma}}_0(t, x) = \varrho\, C(\widetilde{\mu}, \widetilde{\pi}) \varepsilon(\boldsymbol{v}(t, x)), && (t, x) \in [0, T] \times \Omega,\\
&\partial_t \overline{\boldsymbol{\sigma}}_l(t, x) = -\frac{1}{\tau_{\boldsymbol{\sigma},l}} \overline{\boldsymbol{\sigma}}_l(t, x) + \varrho\, C(\widehat{\mu}, \widehat{\pi}) \varepsilon(\boldsymbol{v}(t, x)), \quad l = 1, \dots, L, && (t, x) \in [0, T] \times \Omega.
\end{aligned}$$

Using the zero initial values (2.19d) we can solve these equations to obtain

$$\begin{aligned}
&\overline{\boldsymbol{\sigma}}_0(t, x) = \varrho \int_0^t C(\widetilde{\mu}, \widetilde{\pi}) \varepsilon(\boldsymbol{v}(s, x)) \mathrm{d}s, && (t, x) \in [0, T] \times \Omega,\\
&\overline{\boldsymbol{\sigma}}_l(t, x) = \varrho \int_0^t \exp\Big(\frac{s - t}{\varrho\, \tau_{\boldsymbol{\sigma},l}}\Big) C(\widehat{\mu}, \widehat{\pi}) \varepsilon(\boldsymbol{v}(s, x)) \mathrm{d}s, \quad l = 1, \dots, L, && (t, x) \in [0, T] \times \Omega.
\end{aligned}$$

Summing up and recalling that $\sum_{l=0}^L \overline{\boldsymbol{\sigma}}_l|_\Omega = \mathbf{0}$ gives

$$\mathbf{0} = \int_0^t C\big(\widetilde{\mu} + H(t - s, x)\widehat{\mu}, \widetilde{\pi} + H(t - s, x)\widehat{\pi}\big) \varepsilon(\boldsymbol{v}(s, x)) \mathrm{d}s, \quad (t, x) \in [0, T] \times \Omega, \tag{4.1}$$

where

$$H(t, x) := \sum_{l=1}^L \exp\Big(\frac{-t}{\varrho(x)\, \tau_{\boldsymbol{\sigma},l}}\Big).$$

We take the trace on both sides of (4.1) which results in

$$0 = \int_0^t D(t - s, x) \operatorname{div} \boldsymbol{v}(s, x)\, \mathrm{d}s, \quad (t, x) \in [0, T] \times \Omega, \tag{4.2}$$

with

$$D(t, x) := 3\big(\widetilde{\pi}(x) + H(t, x)\widehat{\pi}(x)\big) - 4\big(\widetilde{\mu}(x) + H(t, x)\widehat{\mu}(x)\big).$$

Next, we differentiate (4.2) with respect to t to get

$$0 = D(0, x) \operatorname{div} \boldsymbol{v}(t, x) + \int_0^t \partial_t D(t - s, x) \operatorname{div} \boldsymbol{v}(s, x)\, \mathrm{d}s, \quad (t, x) \in [0, T] \times \Omega.$$

When $D(0, \cdot) \neq 0$ almost everywhere we have a second kind Volterra integral equation for $\operatorname{div} \boldsymbol{v}(\cdot, x)$ for almost every $x \in \Omega$. The corresponding integral kernel is continuous in (t, s), so we have a unique solution, see, e.g., [35, Thm. 3.10]. Thus, $\operatorname{div} \boldsymbol{v}(t, x) = 0$, $(t, x) \in [0, T] \times \Omega$. Contradicting $\operatorname{div} \boldsymbol{v}(T, \cdot)|_\Omega \not\equiv \mathbf{0}$. So we must have

$$0 = D(0, \cdot) = 3\big(\widetilde{\pi}(\cdot) + L\widehat{\pi}(\cdot)\big) - 4\big(\widetilde{\mu}(\cdot) + L\widehat{\mu}(\cdot)\big) \quad \text{in } \Omega, \tag{4.3}$$

that is, we are in the situation of (4.2) with D replaced by $\partial_t D$. Repeating the above procedure, we conclude that

$$0 = \partial_t D(0,\cdot) = \partial_t H(0,\cdot)\big(3\widehat{\pi}(\cdot) - 4\widehat{\mu}(\cdot)\big) \quad \text{almost everywhere in } \Omega.$$

Since $\partial_t H(0,\cdot) = -\frac{1}{\varrho(\cdot)}\sum_{l=1}^{L}\frac{1}{\tau_{\sigma,l}} \neq 0$ almost everywhere, we must have

$$3\widehat{\pi} = 4\widehat{\mu} \quad \text{and} \quad 3\widetilde{\pi} = 4\widetilde{\mu} \quad \text{in } \Omega \tag{4.4}$$

where the latter equality follows from the former plugged into (4.3).

We proceed by taking the deviatoric (trace-free) part of both sides of (4.1):

$$\mathbf{0} = \int_0^t \big(\widetilde{\mu} + H(t-s,x)\widehat{\mu}\big)\,\mathrm{dev}\,\boldsymbol{v}(s,x)\mathrm{d}s, \quad (t,x) \in [0,T]\times\Omega.$$

This equation is of the same type as (4.2). Accordingly, we get the following Volterra integral equation of the second kind

$$\mathbf{0} = (\widetilde{\mu} + L\widehat{\mu})\,\mathrm{dev}\,\boldsymbol{v}(t,x) + \int_0^t \partial_t H(t-s,x)\widehat{\mu}\;\mathrm{dev}\,\boldsymbol{v}(s,x)\mathrm{d}s, \quad (t,x) \in [0,T]\times\Omega,$$

which, for $\widetilde{\mu}(\cdot)+L\widehat{\mu}(\cdot) \neq 0$ almost everywhere, admits the unique solution $\mathrm{dev}\,\boldsymbol{v}(t,x) = \mathbf{0}$, $(t,x) \in [0,T]\times\Omega$. This contradicts $\mathrm{dev}\,\boldsymbol{v}(T,\cdot)|_\Omega \not\equiv \mathbf{0}$. Hence,

$$\widetilde{\mu} + L\widehat{\mu} = 0 \quad \text{in } \Omega \tag{4.5}$$

and

$$\mathbf{0} = \int_0^t \partial_t H(t-s,x)\widehat{\mu}\;\mathrm{dev}\,\boldsymbol{v}(s,x)\mathrm{d}s, \quad (t,x) \in [0,T]\times\Omega.$$

Differentiating with respect to t and arguing as before leads to $\widehat{\mu} = 0$ in Ω, which, in view of (4.5) and (4.4), yields first $\widehat{\pi} = \widetilde{\pi} = \widehat{\mu} = \widetilde{\mu} = 0$ in Ω and then, by (2.19f) and (2.19e), $h_2 = h_3 = h_4 = h_5 = 0$ in Ω. Finally, the local linear independence (2.15) finishes the proof. □

Local uniqueness of the seismic inverse problem (2.18) follows immediately.

Theorem 4.2. *Suppose that $\boldsymbol{p} = (\varrho, v_\mathrm{S}, \tau_\mathrm{S}, v_\mathrm{P}, \tau_\mathrm{P})^\top \in V_+^5$. Then, $\Phi'(\boldsymbol{p}) \in \mathcal{L}(V^5, \mathcal{W})$ is an injective mapping, and we have that*

$$\min\big\{\|\Phi'(\boldsymbol{p})[\boldsymbol{h}]\|_\mathcal{W} \;:\; \boldsymbol{h} \in V^5,\ \|\boldsymbol{h}\|_{V^5} = 1\big\} > 0.$$

Proof. Assume the minimum to be zero. As V^5 is finite dimensional and $\Phi'(\boldsymbol{p})$ is continuous, there exists an $\boldsymbol{h} \in V^5$ with $\|\boldsymbol{h}\|_{V^5} = 1$ such that $\Phi'(\boldsymbol{p})[\boldsymbol{h}]\boldsymbol{f} = 0$ for all $\boldsymbol{f} \in W_0^{2,1}(\Sigma)$. But then $\boldsymbol{h} = \mathbf{0}$ by Proposition 4.1, which contradicts $\|\boldsymbol{h}\|_{V^5} = 1$. □

In Theorem 2.3 we have seen that the derivative of the FWI forward operator is Lipschitz continuous. Now an application of Lemma C.1 from [19] yields our main result, that is, Lipschitz stability (4.6) of the semi-discrete seismic inverse problem and the TCC (4.7) for the semi-discrete FWI forward operator.

Theorem 4.3. *For any $\boldsymbol{p} = (\varrho, v_\mathrm{S}, \tau_\mathrm{S}, v_\mathrm{P}, \tau_\mathrm{P})^\top \in V_+^5$ there exists an open ball $B_r(\boldsymbol{p}) \subset V_+^5$ such that*

$$\|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{V^5} \lesssim \big\|\Phi(\boldsymbol{p}_1) - \Phi(\boldsymbol{p}_2)\big\|_\mathcal{W} \tag{4.6}$$

and

$$\big\|\Phi(\boldsymbol{p}_1) - \Phi(\boldsymbol{p}_2) - \Phi'(\boldsymbol{p}_2)[\boldsymbol{p}_1 - \boldsymbol{p}_2]\big\|_\mathcal{W} \lesssim \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{V^5}\,\|\Phi(\boldsymbol{p}_1) - \Phi(\boldsymbol{p}_2)\|_\mathcal{W} \tag{4.7}$$

for all $\boldsymbol{p}_i \in B_r(\boldsymbol{p})$*,* $i = 1, 2$*.*

Remark 4.4. *In [17] we considered the elastic regime. There we could prove local injectivity and TCC under weaker smoothness assumptions:* $\varrho \in \mathcal{C}^1(\overline{D})$ *and* $v_{\mathrm{S}}, v_{\mathrm{P}} \in \mathcal{C}^2(\overline{D})$*. Moreover, the observation period* T *had to be only large enough for the shear waves to reach all of* Ω *and not all of* D*, that is,* D *and* $T/2$ *can be replaced by* Ω *and* T *in (3.3), respectively. We conjecture this to hold in the viscoelastic situation as well.*

Remark 4.5. *Let* $\{\boldsymbol{f}_1, \dots, \boldsymbol{f}_m\} \subset W^{2,1}(\Sigma)$ *be the set of* m *sources whose existence is guaranteed by Theorem 3.4 and define a corresponding parameter-to-source map according to*

$$\Theta\colon V_+^5 \subset V^5 \to \mathcal{C}([0,T], X_\Omega)\big)^m, \quad \boldsymbol{p} \mapsto \big(\Phi(\boldsymbol{p})\boldsymbol{f}_1, \dots, \Phi(\boldsymbol{p})\boldsymbol{f}_m\big)^\top,$$

where Φ *is as in (2.17). Then, the local uniqueness result of Theorem 4.2 carries over to* Θ *yielding local Lipschitz stability and TCC as well. Please note that* Θ *models a multi-shot experiment in the language of geophysics. Unfortunately, neither* m *nor the* $\boldsymbol{f}_i$*'s are known explicitly.*

5. Conclusion and discussion

In this work we have extended the results from [19] for FWI in the acoustic regime to the viscoelastic regime: TCC holds for the corresponding forward operator and hence a variety of Newton-like solvers for the seismic inverse problem (2.18) are locally convergent regularization schemes. Moreover, (2.18) is locally wellposed.

From a practical point of view one would like to span the parameter space V by basis functions having a compact support (the local case) and less regularity. The resulting difficulties have already been discussed in Remarks 3.3 and A.3 of [19] which apply to the present situation with straightforward modifications.

Appendix A. Proof of Theorem 2.3

We decompose

$$\Phi = \Psi \circ F \circ B, \tag{A.1}$$

where $B\colon \mathcal{P} \subset L^\infty(D)^5 \to \mathcal{L}^*(X) := \{J \in \mathcal{L}(X) : J^* = J\}$ is given by

$$(\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}}) \mapsto \left(\begin{bmatrix} \boldsymbol{w} \\ \boldsymbol{\psi}_0 \\ \vdots \\ \boldsymbol{\psi}_L \end{bmatrix} \mapsto \begin{bmatrix} \varrho \boldsymbol{w} \\ C^{-1}(\mu, \pi)\boldsymbol{\psi}_0 \\ C^{-1}(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\boldsymbol{\psi}_1 \\ \vdots \\ C^{-1}(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\boldsymbol{\psi}_L \end{bmatrix} \right)$$

where μ and π have been defined in (2.9).

Furthermore, we define the mapping

$$F\colon \mathsf{D}(F) \subset \mathcal{L}^*(X) \to \underbrace{\mathcal{L}\big(W_0^{2,1}(D), \mathcal{C}([0,T], X)\big)}_{=:\, \mathcal{S}}, \tag{A.2}$$

$$P \mapsto \big(\boldsymbol{f} \mapsto u = (\boldsymbol{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L)^\top\big),$$

where u is the classical solution of (2.6) with $f = (\boldsymbol{f}, \boldsymbol{0}, \dots, \boldsymbol{0})^\top$, $u_0 = 0$, A from (2.5), and B is replaced by

$$P \in \mathsf{D}(F) := \big\{\Lambda \in \mathcal{L}^*(X) \; : \; \lambda_- \|x\|_X^2 < \langle \Lambda x, x\rangle_X < \lambda_+ \|x\|_X^2\big\}$$

for some $0 < \lambda_- < \lambda_+ < \infty$. In view of (2.8) and (2.12) we can choose λ_- and λ_+ depending only on the constants defining $\mathcal{P}$ such that $B(\mathcal{P}) \subset \mathsf{D}(F)$. Then, the factorization of Φ in (A.1) is well defined and we obtain that

$$\Phi'(\boldsymbol{p}) = \Psi F'(B(\boldsymbol{p}))B'(\boldsymbol{p}), \qquad \boldsymbol{p} = (\varrho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}}) \in \mathcal{P}.$$

The differentiability of F is established in the following theorem and B' is explicitly given in (A.5) below.

Theorem A.1. *The map F defined in (A.2) is Fréchet-differentiable at $P \in \mathsf{D}(F)$ where*

$$F'(P)[H]f = \overline{u} \quad \text{for } H \in \mathcal{L}^*(X)$$

with $\overline{u} \in \mathcal{C}\big([0,T], X\big)$ being the mild (in fact the classical) solution of

$$P\overline{u}'(t) + (A + PQ)\overline{u}(t) = -H(u'(t) + Qu(t)),\ t \in [0,T], \quad \overline{u}(0) = 0,$$

where $u = F(P)f$. Moreover, F' is Lipschitz continuous, that is,

$$\|F'(P_1) - F'(P_2)\|_{\mathcal{L}(\mathcal{L}^*(X),\mathcal{S})} \lesssim \|P_1 - P_2\|_{\mathcal{L}(X)}, \qquad P_1, P_2 \in \mathsf{D}(F), \tag{A.3}$$

where the Lipschitz constant only depends on T, λ_-, and λ_+.

Proof. The Fréchet-differentiability is already shown in [34, Thm. 3.2]. The Lipschitz continuity of F' follows by a slight modification of the proof of Theorem B.2 from [19] in which a similar result is formulated for $Q = 0$ in (2.6). □

For any $\boldsymbol{p_i} \in \mathcal{P}$, $i = 1, 2$, using (A.3) we proceed with

$$\begin{aligned}
\|\Phi'(\boldsymbol{p}_1) - \Phi'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{W})} &\le \|F'(B(\boldsymbol{p}_1))B'(\boldsymbol{p}_1) - F'(B(\boldsymbol{p}_2))B'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{S})} \\
&\le \big\|\big(F'(B(\boldsymbol{p}_1)) - F'(B(\boldsymbol{p}_2))\big)B'(\boldsymbol{p}_1)\big\|_{\mathcal{L}(V^5,\mathcal{S})} \\
&\quad + \big\|F'(B(\boldsymbol{p}_2))\big(B'(\boldsymbol{p}_1) - B'(\boldsymbol{p}_2)\big)\big\|_{\mathcal{L}(V^5,\mathcal{S})} \\
&\lesssim \|B(\boldsymbol{p}_1) - B(\boldsymbol{p}_2)\|_{\mathcal{L}(X)}\|B'(\boldsymbol{p}_1)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))} \\
&\quad + \|F'(B(\boldsymbol{p}_2))\|_{\mathcal{L}(\mathcal{L}^*(X),\mathcal{S})}\|B'(\boldsymbol{p}_1) - B'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))}.
\end{aligned} \tag{A.4}$$

As $\mathcal{P}$ is bounded, the Lipschitz continuity (A.3) further yields that $\|F'(B(\boldsymbol{p}))\|_{\mathcal{L}(\mathcal{L}^*(X),\mathcal{S})} \lesssim 1$ uniformly in $\boldsymbol{p} \in \mathcal{P}$. In the next step we show uniform boundedness of $\|B'(\cdot)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))}$ in $\mathcal{P}$. From [34, Sec. 4.2] we have, for $\boldsymbol{h} = (h_1, h_2, h_3, h_4, h_5) \in V^5$, that

$$B'(\boldsymbol{p})\boldsymbol{h}\begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{\psi}_0 \\ \vdots \\ \boldsymbol{\psi}_L \end{pmatrix} = \begin{pmatrix} h_1\,\boldsymbol{w} \\ -\frac{h_1}{\varrho}\,C^{-1}(\mu,\pi)\boldsymbol{\psi}_0 + \varrho\,(C^{-1})'(\mu,\pi)\begin{bmatrix}\widetilde{\mu} \\ \widetilde{\pi}\end{bmatrix}\boldsymbol{\psi}_0 \\ -\frac{h_1}{\varrho}\,C^{-1}(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\boldsymbol{\psi}_1 + \varrho\,(C^{-1})'(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\begin{bmatrix}\widehat{\mu} \\ \widehat{\pi}\end{bmatrix}\boldsymbol{\psi}_1 \\ \vdots \\ -\frac{h_1}{\varrho}\,C^{-1}(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\boldsymbol{\psi}_L + \varrho\,(C^{-1})'(\tau_{\mathrm{S}}\mu, \tau_{\mathrm{P}}\pi)\begin{bmatrix}\widehat{\mu} \\ \widehat{\pi}\end{bmatrix}\boldsymbol{\psi}_L \end{pmatrix} \tag{A.5}$$

where $\widetilde{\mu}$, $\widetilde{\pi}$, and $\widehat{\mu}$, $\widetilde{\pi}$, are linear functions of $h_2, \ldots, h_5$, see (2.19e) and (2.19f), respectively. Further,

$$(C^{-1})'(\mu,\pi)\begin{bmatrix}\widetilde{\mu} \\ \widetilde{\pi}\end{bmatrix} = -C(\mu,\pi)^{-1}C(\widetilde{\mu},\widetilde{\pi})C(\mu,\pi)^{-1}.$$

This explicit expression of B' together with (2.8) implies $\|B'(\boldsymbol{p})\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))} \lesssim 1$ for $\boldsymbol{p} \in \mathcal{P}$. Substituting these estimates into (A.4) we obtain that

$$\begin{aligned}\|\Phi'(\boldsymbol{p}_1) - \Phi'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{W})} &\lesssim \|B(\boldsymbol{p}_1) - B(\boldsymbol{p}_2)\|_{\mathcal{L}(X)} + \|B'(\boldsymbol{p}_1) - B'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))}\\ &\lesssim \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{V^5} + \|B'(\boldsymbol{p}_1) - B'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))},\end{aligned}$$

where we used the mean value theorem and the estimate $\|B'(\boldsymbol{p})\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))} \lesssim 1$ for any $\boldsymbol{p} \in \mathcal{P}$ in the last step. The final estimate

$$\|B'(\boldsymbol{p}_1) - B'(\boldsymbol{p}_2)\|_{\mathcal{L}(V^5,\mathcal{L}^*(X))} \lesssim \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|_{V^5}$$

can be validated either directly using (A.5) or by applying again the mean value theorem together with the estimate $\|B''(\boldsymbol{p})[\boldsymbol{h}_1, \boldsymbol{h}_2]\|_{\mathcal{L}(X)} \lesssim \|\boldsymbol{h}_1\|_{V^5}\|\boldsymbol{h}_2\|_{V^5}$ for $\boldsymbol{p} \in \mathcal{P}$ and $\boldsymbol{h}_1, \boldsymbol{h}_2 \in V^5$, where the second derivative $B''(\boldsymbol{p})$ is given in [34, Sec 4.3].

References

[1] G. Alessandrini, M. V. de Hoop, F. Faucher, R. Gaburro, and E. Sincich, *Inverse problem for the Helmholtz equation with Cauchy data: reconstruction with conditional well-posedness driven iterative regularization*, ESAIM Math. Model. Numer. Anal., 53 (2019), pp. 1005–1030, https://doi.org/10.1051/m2an/2019009.

[2] G. Alessandrini and S. Vessella, *Lipschitz stability for the inverse conductivity problem*, Adv. in Appl. Math., 35 (2005), pp. 207–241, https://doi.org/10.1016/j.aam.2004.12.002.

[3] E. Beretta, M. V. de Hoop, F. Faucher, and O. Scherzer, *Inverse boundary value problem for the Helmholtz equation: quantitative conditional Lipschitz stability estimates*, SIAM J. Math. Anal., 48 (2016), pp. 3962–3983, https://doi.org/10.1137/15M1043856.

[4] E. Beretta, M. V. de Hoop, E. Francini, S. Vessella, and J. Zhai, *Uniqueness and Lipschitz stability of an inverse boundary value problem for time-harmonic elastic waves*, Inverse Problems, 33 (2017), pp. 035013, 27, https://doi.org/10.1088/1361-6420/aa5bef.

[5] E. Beretta, M. V. de Hoop, and L. Qiu, *Lipschitz stability of an inverse boundary value problem for a Schrödinger-type equation*, SIAM J. Math. Anal., 45 (2013), pp. 679–699, https://doi.org/10.1137/120869201.

[6] J. O. Blanch, J. O. A. Robertsson, and W. W. Symes, *Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique*, Geophysics, 60 (1995), pp. 176–184, https://doi.org/10.1190/1.1443744, https://doi.org/10.1190/1.1443744.

[7] T. Bohlen, *Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme*, PhD thesis, Christian-Albrechts-Universität zu Kiel, 1998, https://bit.ly/2LMOSWr.

[8] T. Bohlen, *Parallel 3-D viscoelastic finite difference seismic modelling*, Comput. Geosci., 28 (2002), pp. 887–899, https://doi.org/10.1016/S0098-3004(02)00006-7.

[9] T. Bohlen, M. R. Fernandez, J. Ernesti, C. Rheinbay, A. Rieder, and C. Wieners, *Visco-acoustic full waveform inversion: From a DG forward solver to a Newton-CG inverse solver*, Comput. Math. Appl., 100 (2021), pp. 126–140, https://doi.org/10.1016/j.camwa.2021.09.001.

[10] P. G. Ciarlet, *Mathematical elasticity. Vol. I*, vol. 20 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.

[11] C. M. Dafermos, *An abstract Volterra equation with applications to linear viscoelasticity*, J. Differential Equations, 7 (1970), pp. 554–569, https://doi.org/10.1016/0022-0396(70)90101-4.

[12] M. V. de Hoop, C.-L. Lin, and G. Nakamura, *Holmgren-John unique continuation theorem for viscoelastic systems*, in Time-dependent Problems in Imaging and Parameter Identification [32], pp. 287–301, https://doi.org/10.1007/978-3-030-57784-1_10.

[13] M. V. de Hoop, L. Qiu, and O. Scherzer, *Local analysis of inverse problems: Hölder stability and iterative reconstruction*, Inverse Problems, 28 (2012), pp. 045001, 16, https://doi.org/10.1088/0266-5611/28/4/045001.

[14] W. Dörfler and C. Wieners, *Space-time approximations for linear acoustic, elastic, and electromagnetic wave equations*, in Wave Phenomena. Mathematical Analysis and Numerical Approximation, vol. 49 of Oberwolfach Seminars, Birkhäuser Cham, 2023, https://link.springer.com/book/9783031057946.

[15] G. Duvaut and J.-L. Lions, *Inequalities in mechanics and physics*, vol. 219 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin-New York, 1976. Translated from the French by C. W. John.

[16] H. Egger and B. Hofmann, *Tikhonov regularization in Hilbert scales under conditional stability assumptions*, Inverse Problems, 34 (2018), pp. 115015, 17, https://doi.org/10.1088/1361-6420/aadef4.

[17] M. Eller, R. Griesmaier, and A. Rieder, *Tangential cone condition for the full waveform forward operator in the elastic regime: the non-local case*, CRC 1173 Preprint 2022/48, Karlsruhe Institute of Technology, 2022, https://doi.org/10.5445/IR/1000150845.

[18] M. Eller, N. Honda, C.-L. Lin, and G. Nakamura, *Global unique continuation from the boundary for a system of viscoelasticity with analytic coefficients and a memory term*, Inverse Probl. Imaging, 16 (2022), pp. 1529–1542, https://doi.org/10.3934/ipi.2022049.

[19] M. Eller and A. Rieder, *Tangential cone condition and Lipschitz stability for the full waveform forward operator in the acoustic regime*, Inverse Problems, 37 (2021), pp. Paper No. 085011, 17, https://doi.org/10.1088/1361-6420/ac11c5.

[20] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1996.

[21] I. Epanomeritakis, V. Akçelik, O. Ghattas, and J. Bielak, *A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion*, Inverse Problems, 24 (2008), pp. 034015, 26, https://doi.org/10.1088/0266-5611/24/3/034015.

[22] A. Fichtner, *Full seismic waveform modelling and inversion*, Advances in Geophysical and Environmental Mechanics and Mathematics, Springer-Verlag Berlin Heidelberg, 2011, https://doi.org/10.1007/978-3-642-15807-0.

[23] M. Hanke, *Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems*, Numer. Funct. Anal. Optim., 18 (1997), pp. 971–993, https://doi.org/10.1080/01630569708816804.

[24] M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), pp. 21–37, https://doi.org/10.1007/s002110050158.

[25] B. Harrach, *Uniqueness and Lipschitz stability in electrical impedance tomography with finitely many electrodes*, Inverse Problems, 35 (2019), pp. 024005, 19, https://doi.org/10.1088/1361-6420/aaf6fc.

[26] K. Ito and F. Kappel, *Evolution equations and approximations*, vol. 61 of Series on Advances in Mathematics for Applied Sciences, World Scientific Publishing Co., Inc., River Edge, NJ, 2002, https://doi.org/10.1142/9789812777294.

[27] B. Kaltenbacher, *Regularization based on all-at-once formulations for inverse problems*, SIAM J. Numer. Anal., 54 (2016), pp. 2594–2618, https://doi.org/10.1137/16M1060984.

[28] B. Kaltenbacher and M. L. P. de Souza, *Convergence and adaptive discretization of the IRGNM Tikhonov and the IRGNM Ivanov method under a tangential cone condition in Banach space*, Numer. Math., 140 (2018), pp. 449–478, https://doi.org/10.1007/s00211-018-0971-5.

[29] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, vol. 6 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008, http://dx.doi.org/10.1515/9783110208276.

[30] B. Kaltenbacher, T. T. N. Nguyen, and O. Scherzer, *The tangential cone condition for some coefficient identification model problems in parabolic pdes*, in Time-dependent Problems in Imaging and Parameter Identification [32], pp. 121–163, https://doi.org/10.1007/978-3-030-57784-1_5.

[31] B. Kaltenbacher, F. Schöpfer, and T. Schuster, *Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems*, Inverse Problems, 25 (2009), pp. 065003, 19, https://doi.org/10.1088/0266-5611/25/6/065003.

[32] B. Kaltenbacher, T. Schuster, and A. Wald, *Time-dependent Problems in Imaging and Parameter Identification*, Springer, Cham, 2021, https://doi.org/10.1007/978-3-030-57784-1.

[33] A. Kirsch and A. Rieder, *Inverse problems for abstract evolution equations with applications in electrodynamics and elasticity*, Inverse Problems, 32 (2016), pp. 085001, 24, https://doi.org/10.1088/0266-5611/32/8/085001.

[34] A. Kirsch and A. Rieder, *Inverse problems for abstract evolution equations II: Higher order differentiability for viscoelasticity*, SIAM J. Appl. Math., 79 (2019), pp. 2639–2662, https://doi.org/10.1137/19M1269403. Corrigendum under https://doi.org/10.48550/arXiv.2203.01309.

[35] R. Kress, *Linear integral equations*, vol. 82 of Applied Mathematical Sciences, Springer, New York, third ed., 2014, https://doi.org/10.1007/978-1-4614-9593-2.

[36] A. Lechleiter and A. Rieder, *Newton regularizations for impedance tomography: convergence by local injectivity*, Inverse Problems, 24 (2008), pp. 065009, 18, https://doi.org/10.1088/0266-5611/24/6/065009.

[37] L. Métivier, R. Brossier, J. Virieux, and S. Operto, *Full waveform inversion and the truncated Newton method*, SIAM J. Sci. Comput., 35 (2013), pp. B401–B437, https://doi.org/10.1137/120877854.

[38] L. Pieronek and A. Rieder, *On the iterative regularization of non-linear illposed problems in L^∞*, CRC 1173 Preprint 2021/46, Karlsruhe Institute of Technology, dec 2021, https://www.waves.kit.edu/downloads/CRC1173_Preprint_2021-46.pdf.

[39] A. Rieder, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal., 43 (2005), pp. 604–622, https://doi.org/10.1137/040604029.

[40] J. O. Robertsson, J. O. Blanch, and W. W. Symes, *Viscoelastic finite-difference modeling*, Geophysics, 59 (1994), pp. 1444–1456, http://dx.doi.org/10.1190/1.1443701.

[41] O. Scherzer, *The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems*, Computing, 51 (1993), pp. 45–60, https://doi.org/10.1007/BF02243828.

[42] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski, *Regularization methods in Banach spaces*, vol. 10 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2012, https://doi.org/10.1515/9783110255720.

[43] A. Tarantola, *Theoretical background for the inversion of seismic waveforms including elasticity and attenuation*, Pure and Applied Geophysics, 128 (1988), pp. 365–399, https://doi.org/10.1007/BF01772605.

[44] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, Geophysics, 74 (2009), pp. 1–26, https://doi.org/10.1190/1.3238367.

[45] P. Yang, R. Brossier, L. Métivier, J. Virieux, and W. Zhou, *A time-domain preconditioned truncated Newton approach to visco-acoustic multiparameter full waveform inversion*, SIAM J. Sci. Comput., 40 (2018), pp. B1101–B1130, https://doi.org/10.1137/17M1126126.

[46] U. Zeltmann, *The viscoelastic seismic model: existence, uniqueness and differentiability with respect to parameters*, PhD thesis, Karlsruhe Institute of Technology, 2018, http://dx.doi.org/10.5445/IR/1000093989.

Department of Mathematics, Georgetown University, Washington, DC 20057, USA
Email address: mme4@georgetown.edu

Department of Mathematics, Karlsruhe Institute of Technology, D-76128 Karlsruhe, Germany
Email address: roland.griesmaier@kit.edu
Email address: andreas.rieder@kit.edu